前　言

在数学上，当人们在若干个条件下证明了某一结论时，常希望知道：所证明的结论对所设条件依赖到何种程度．具体地说，设A是条件之一，当免除条件A时，定理的结论是否仍成立．或者，人们希望知道：由某些特定的条件出发能否证明某项结论．考察这个问题的一个方法是设法构造反例．常见的反例有两种：其一是当从一组条件中免除条件A时，定理的结论不再成立．另一种是令某些条件满足，但某项特定的断言仍不真．构造反例不仅是使理论变得更加完整，成为不可缺少的一环，而且对于深入理解正面命题的内容、证明的思想和技巧以及定理条件在证明中的作用都有着极其重要的意义．

概率论与数理统计是研究大量随机现象规律性的学科．现在高等院校理、工、医、农、财经的许多专业都设置了概率论与数理统计课程，有的学校还设立了数理统计系或统计运筹系．概率论与数理统计的理论和方法已广泛地应用于工业、农业、国防和科学技术中，成为数学中活跃的分支之一．

近年来，国内已出版了许多内容深浅不一、风格不同的概率统计教材和参考书，但是在国内外出版的许多概率统计教本或专著中，还没见到一本较为全面、系统地论述概率统计中反例的著作．本书的写作是希望能多少弥补这个缺陷．

本书针对概率论与数理统计的基本概念和正面命题，较为

全面、系统地提供反例并给予证明．全书共收集反例一百三十七条，内容涉及概率空间、随机变量、分布函数、独立性（含鞅论）、数字特征、母函数、特征函数（含无穷可分分布）、极限理论、完全统计量与充分统计量、参数点估计、假设检验、线性模型、抽样理论等．每一条反例，大致由四个部分的内容组成：(1) 需要用到的有关定义和定理；(2) 正面命题；(3) 引用或构造反例并给予证明（太复杂的只提供文献）.这部分是重点；(4) 进一步的讨论（介绍文献和专著中更深入的结果）．对于内容比较深入的反例，我们在该例前面标有(*) 号，而且对阅读这样的反例所需用到的知识，我们均不作详细介绍，因此希望还不具备这方面知识的读者，初次阅读时可以略去，待必要时再参考所列文献或专著补看．

本书概率论部分我们用ξ，η，ξ……或ξ_1，ξ_2，…，ξ_n等等表示随机变量，而在数理统计部分，我们则用X，Y，Z，……或X_1，X_2，…，X_n表示随机变量．每个反例前面的数字，如6—13，表示第六章的第13个反例，余类推．另外定理、定义、例子、注记的编号都按各个反例单独进行．

我们期望本书对于具有一般概率统计知识的读者的进一步提高能有所裨益．

本书在写作过程中，曾得到中国科学技术大学陈希孺教授的热情鼓励和指导．初稿完成后，又承他和安徽大学的陈桂景同志仔细审阅并提出许多宝贵意见，作者谨向他们表示衷心的感谢．

限于作者的水平，书中存在的问题肯定不少，作者恳切地希望专家和读者提出宝贵的批评意见．

张尚志　刘锦萼

1984.11

目　　录

第一章　与随机事件、概率空间、古典概型有关的反例（1）

1—1　基本事件但不是事件的例子……………………（1）

1—2　同一随机现象可以用不同的样本空间来描述的例子…………………………………………（2）

1—3　两事件互斥但不互逆的例子…………………（4）

1—4　概率为0的事件并非是不可能事件的例子……………………………………………………（5）

1—5　概率为1的事件并非是必然事件的例子……（6）

(*)1—6　集类E上的非负集函数μ具有有限可加性、连续性但不具有可列可加性的例子……………（6）

1—7　上极限事件和下极限事件不相等的例子……（11）

1—8　非古典型随机试验的例子……………………（12）

第二章　与独立性、条件概率有关的反例………………（14）

2—1　两事件不独立的例子…………………………（14）

2—2　两事件不独立，但不一定互斥的例子………（16）

2—3　事件两两独立，但并非相互独立的例子……（17）

2—4　$P(ABC)=P(A)P(B)P(C)$，但A，B，C三事件并非两两独立的例子…………………（19）

2—5　说明$P(A)$，$P(A|B)$，$P(AB)$三者不同含义的例子……………………………………………（22）

2—6 非独立试验概型（非Bernoulli 概型）的例子 ……（24）

(*)2—7 $\varlimsup_{n\to\infty}\sup E|\xi_n|<\infty$，但 $\xi_n\xrightarrow{L_1}\xi$不成立的例子 ……（25）

(*)2—8 并非任何一个鞅$\{\xi_n\}$均存在一个满足$E|\xi|<\infty$的随机变量ξ和概率空间（Ω，$\mathscr{F}_1$，P）中$\mathscr{F}_1$的子σ代数$\mathscr{D}_n$使得$\xi_n=E\{\xi|\mathscr{D}_n\}$的例子 ……（28）

(*)2—9 ξ与ζ独立，但$E(\xi|\eta,\zeta)\neq E(\xi|\eta)$的例子 ……（29）

第三章 与随机变量，分布函数有关的反例 ……（31）

3—1 $\xi(\omega)$是定义在概率空间（Ω，$\mathscr{F}$，P）上的单值实函数，但它不是（Ω，$\mathscr{F}_1$，P）上的随机变量的例子 ……（31）

(*)3—2 随机变量的勒贝格可测函数不一定是随机变量的例子 ……（32）

3—3 既非离散型，又非连续型的分布函数的例子 ……（37）

3—4 设ξ是一个连续型随机变量，g是某个连续函数，$\eta=g(\xi)$不是连续型随机变量的例子 ……（41）

3—5 不同的随机变量（向量），具有相同的分布函数的例子 ……（43）

3—6 连续型随机变量之密度函数未必是连续的例子 ……（44）

3—7 二元函数$F(x,y)$对每个变元非降，左连续且$F(x,-\infty)=0$，$F(-\infty,y)=0$，$F(+\infty,$

$+\infty)=1$但仍不是分布函数的例子…………(46)
3—8 边际分布是正态分布，但联合分布不是多元正态分布的例子……………………………(48)
3—9 随机变量ξ、η相互独立，而且同分布，但不一定有 $\xi=\eta(\alpha,s)$ 成立的例子………………(53)
3—10 ξ、η 同分布但不独立时，$\xi=\xi-\eta$ 不一定是对称随机变量的例子…………………………(54)
(*)3—11 ξ、η服从正态分布，但ξ、η不独立，则$\xi+\eta$不一定服从正态分布的例子…………………(56)
3—12 两个不同的联合分布（函数），它们可以有相同的边际分布的例子…………………………(58)
3—13 随机变量ξ_1，ξ_2，ξ_3，两两独立，但不相互独立的例子……………………………………(62)
3—14 ξ、η不独立，但ξ^2和η^2独立的例子…………(66)
3—15 相同的随机向量构造的不同的 Borel 可测函数（不恒等于常数）之间也可能是独立的例子。…………………………………………(69)
3—16 从随机向量$\xi=(\xi_1, \xi_2, \cdots, \xi_m)$和$\eta=(\eta_1, \eta_2, \cdots, \eta_n)$之间的独立性推不出$\xi$的分量$\xi_1, \xi_2, \cdots, \xi_m$或者$\eta$的分量$\eta_1, \eta_2, \cdots, \eta_n$之间的独立性的例子……………………………………(72)
3—17 ξ和η_1独立，ξ和η_2独立，但ξ和随机向量(η_1, η_2)不独立的例子……………………………(74)
(*)3—18 非独立随机变量序列的例子…………………(76)
(*)3—19 分布函数F_1和F_2的卷积绝对连续，但F_1和F_2不绝对连续的例子………………………………(77)
(*)3—20 随机变量ξ、η的各阶矩不全存在，即使ξ、η

已经存在的各阶矩相等，亦不能推出ξ，η的分布相同的例子…………………………（78）
(*)3—21 ξ与$\frac{1}{\xi}$有相同的分布，而ξ并非服从哥西分布的例子…………………………（79）
(*)3—22 ξ与$1-\xi$有相同的分布，而ξ并非服从$B(\alpha,\alpha)$分布的例子…………………………（81）
(*)3—23 ξ与η独立同分布，$\zeta=\frac{\xi}{\eta}\sim\mathscr{C}(1,0)=\frac{1}{\pi(1+x^2)}$，但$\zeta$，$\eta$并非服从正态分布的例子…………………………（83）
(*)3—24 ξ与η独立，分别服从$B(\alpha_1,\beta_1)$及$B(\alpha_2,\beta_2)$分布，又$\xi\eta\sim B(\alpha,\beta)$，则$\beta=\beta_1+\beta_2$，但只有$\alpha=\alpha_1$或$\alpha=\alpha_2$的例子…………………………（85）
(*)3—25 $\sum_{k=1}^{\infty}\frac{m_k}{k!}S^k$ ($m_k=E\xi^k$) 对S不绝对收敛，$\{m_k\}$仍唯一决定ξ之分布函数的例子…………………………（85）
第四章 与数字特征有关的反例…………………………（87）
4—1 随机变量的数学期望不存在（从而方差也不存在）的例子…………………………（87）
4—2 随机变量的数学期望存在，但方差不存在的例子…………………………（89）
4—3 任何阶矩都不存在的随机变量的例子…………（92）
4—4 随机变量ξ的一阶矩存在，但没有更高整数阶矩的例子…………………………（95）
4—5 随机变量ξ在$0<r<1$时，$E\xi^r$存在；但在$r\geq1$时，$E\xi^r$不存在的例子…………………………（96）

4—6 随机变量ξ_1和ξ_2，它们的一切整数阶矩都相同（即$E\xi_1^k=E\xi_2^k$，$k=1,2,3\cdots$），但它们的分布函数不相等的例子……………………(97)
4—7 随机变量ξ_1和ξ_2不相关，但也不独立的例子……………………(101)
4—8 相关系数$\rho(\xi_1,\xi_2)>0$，$\rho(\xi_2,\xi_3)>0$，但$\rho(\zeta_1,\zeta_3)<0$的例子……………………(105)
4—9 ξ_1和ξ_2不独立，但$E(\xi_1\cdot\xi_2)=E\xi_1\cdot E\xi_2$的例子……………………(106)
4—10 ξ_1,ξ_2独立，但$E(\xi_1+\xi_2)^k\neq E\xi_1^k+E\xi_2^k$的例子（正整数$k\neq1$）……………………(108)
4—11 $E[E(\eta|\xi)]$存在，但$E\eta$不存在，因而$E\eta=E[E(\eta|\zeta)]$不成立的例子……………………(109)
4—12 中位数不唯一的例子……………………(111)
4—13 数学期望不存在，但中位数存在的随机变量的例子……………………(112)
4—14 随机变量的众数不唯一的例子……………………(113)
第五章 与特征函数、母函数有关的反例……………………(114)
5—1 随机变量ξ_1,ξ_2不独立，但$\varphi_{\xi_1+\xi_2}(t)=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t)$成立的例子……………………(114)
5—2 当k为奇数时，随机变量ξ的特征函数$\varphi(t)$在$t=0$处可微分k次，但$E\xi^k$不存在的例子……(122)
5—3 分布函数绝对连续，但其对应的特征函数不绝对可积的例子……………………(125)
5—4 特征函数$\varphi(t)$在有限区间内的值不足以唯一确定此$\varphi(t)$，从而也不足以唯一决定分布函数$F(x)$的例子……………………(126)

5—5 特征函数列的极限函数不是特征函数的例子 ……(129)
5—6 分布函数不具有再生性的例子……(130)
5—7 分布函数$F(x)$不是无穷可分分布的例子……(133)
5—8 无处为 0 的特征函数不是无穷可分的例子 ……(134)
5—9 无穷可分的特征函数可以分解为不是无穷可分的特征函数的乘积的例子……(134)
5—10 随机变量的矩母函数不存在的例子……(137)
5—11 随机变量ξ的各阶矩都存在,但矩母函数不存在的例子……(138)
(*)5—12 并非所有的特征函数都是解析的例子……(139)
(*)5—13 消去法对一般特征函数之分解不一定成立的例子……(140)
(*)5—14 无穷可分的特征函数存在不可分解因子的例子……(140)
第六章 与收敛性有关的反例……(143)
6—1 分布函数列$\{F_n(x)\}$弱收敛于$F(x)$，但$F(x)$不是分布函数的例子……(143)
6—2 分布函数列$F_n(x)\xrightarrow{W}F(x)$，但$F(x)$不唯一的例子……(144)
6—3 即使$\lim\limits_{n\to\infty}\xi_n(\omega)=\xi(\omega)$对$\forall\omega\in\Omega$都成立，也不能保证$\lim\limits_{n\to\infty}F_n(x)=F(x)$对$\forall x\in R_1$成立的例子……(145)
6—4 $F_n(x)\xrightarrow{W}F(x)$，但$E\xi_n^k\to E\xi^k$不成立的例子

……………………………………………………(147)

6—5　$F_n(x)\xrightarrow{W}F(x)$，但$F_n(x)\xrightarrow{C}F(x)$ 不成立的例子……………………………………………(148)

6—6　设分布函数列 $\{F_n(x)\}$ 对应的特征函数列为 $\{\varphi_n(t)\}$，若$\{\varphi_n(t)\}$收敛于某函数$\varphi(t)$，但 $\varphi(t)$在$t=0$ 不连续，则推不出 $\{F_n(x)\}$弱收敛于某分布函数的例子……………………(150)

6—7　由$\xi_n\xrightarrow{L}\xi$ 推不出相应的分布密度函数或概率分布的收敛性的例子…………………………(151)

6—8　$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$，但$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$不成立的例子……………………………………………(153)

6—9　$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$但$E\xi_n^k\to E\xi^k(\forall k\geqslant 1)$ 不成立的例子……………………………………………(155)

6—10　$\xi_n\xrightarrow{P}\xi$ 但$\xi_n\xrightarrow{a,s}\xi$不成立的例子…………………(157)

6—11　波雷尔——康特立引理（1）的逆不成立的例子……………………………………………(161)

6—12　$\xi_n\xrightarrow{P}\xi$但 $\xi_n\xrightarrow{r}\xi$不成立的例子…………………(162)

6—13　设$0<s<r$，$\xi_n\xrightarrow{S}\xi$但 $\xi_n\xrightarrow{r}\xi$ 不成立的例子……………………………………………(163)

6—14　有关 r—阶收敛与几乎处处收敛之间关系的反例……………………………………………(165)

(*)6—15　当$F_n\xrightarrow{C}F$ 时，$\int_{-\infty}^{\infty}gdF_n\nrightarrow\int_{-\infty}^{\infty}gdF$的例子……(167)

(*)6—16　ξ_n依分布收敛到ξ（记为$\xi_n\xrightarrow{L}\xi$），但ξ_n之矩母

函数$M_n(s)$不收敛到ξ之矩母函数 $M(s)$的例子……(168)
(*)6—17 矩母函数的极限函数不是矩母函数的例子……(169)
第七章 与大数定律、中心极限定理有关的反例……(170)
7—1 随机变量序列$\{\xi_n\}$不服从于大数定律的例子……(170)
7—2 随机变量序列$\{\xi_k\}$不满足马尔科夫条件，但服从大数定律的例子……(175)
7—3 独立随机变量序列$\{\xi_k\}$不满足车贝谢夫大数定律的条件，但满足马尔科夫大数定律条件的例子……(178)
7—4 独立随机变量序列$\{\xi_k\}$不满足格涅坚科大数定律的充要条件的例子……(179)
7—5 马尔科夫条件满足，但柯尔莫哥洛夫强大数定律条件不满足的例子……(181)
7—6 独立随机变量序列$\{\xi_k\}$不满足强大数定律的例子……(184)
7—7 林德贝格条件不满足，但中心极限定理仍成立的例子……(187)
7—8 费勒条件不满足，但中心极限定理仍成立的例子……(189)
7—9 有关大数定律和中心极限定理之间关系的反例……(191)
7—10 随机变量序列$\{\xi_n\}$独立同分布，但不服从中心极限定理的例子……(196)
*)7—11 $E|\xi|^2=\infty$，Lindeberg-Iévy 中心极限定理

仍成立的例子……………………………………(197)

(*)7—12 $\{\xi_{nk}1\leqslant k\leqslant n,\ n\geqslant 1\}$满足$u,a,n$条件，但$\max\limits_{1\leqslant k\leqslant n}|\xi_{nk}|\overset{P}{\nrightarrow}0$ 的例子……………………………(198)

(*)7—13 $\sum\limits_{n=1}^{\infty}\frac{D\xi_n}{n^2}\quad\infty$，$\{\xi_n\}$ 仍服从强大数定律的例子……………………………………………(199)

(*)7—14 $\sum\limits_{n=1}^{\infty}(\xi_n-E\xi_n)a,s$ 收敛，但 $\sum\limits_{n=1}^{\infty}D\xi_n=\infty$ 的例子……………………………………………(200)

第八章 与充分统计量和完全统计量有关的反例………(203)

8—1 并非一切统计量都是充分统计量的例子……(204)

(*)8—2 极小充分统计量不是完全充分统计量的例子…………………………………………………(206)

8—3 次序统计量不是完全统计量的例子…………(209)

8—4 有界完全的分布族（或统计量）不是完全的分布族（或统计量）的例子…………………(211)

(*)8—5 可测函数 $f(X)$ 与有界充分完全统计量$t(X)$独立，但$f(X)$的分布与θ 有关的例子………(215)

8—6 充分统计量的函数不是充分统计量的例子…………………………………………………(216)

8—7 充分完全统计量的函数不是充分完全统计量的例子…………………………………………(217)

8—8 指数族分布表达式中的 $T(X)$ 不是充分完全统计量的例子……………………………(217)

(*)8—9 在适当的条件下，当$T_i(X_i)$是 $\theta_i(i=1,2)$的充分统计量时，$(T_1(X_1),\ T_2(X_2))$是 $(\theta_1,$

θ_2）的联合充分统计量的例子……………… (221)

第九章　与点估计有关的反例……………………………… (223)

9—1　参数不存在无偏估计的例子………………… (223)

9—2　无偏估计不是一致最小方差无偏估计的例子
……………………………………………………… (225)

9—3　$X\sim\{f(x;\theta),\ \theta\in\Theta\},x_1,\cdots,x_n$是其$iid$样本，参数 θ 的无偏估计是$X_1,X_2,\cdots,X_n$的对称函数，但不是θ的$UMVUE$的例子……………… (229)

(*)9—4　无偏估计存在，而 $UMVUE$不存在的例子
……………………………………………………… (231)

9—5　$\hat{\theta}$是θ的无偏估计，而$g(\hat{\theta})$不是$g(\theta)$的无偏估计的例子（其中$g(X)$是 $X=(X_1,\cdots,X_n)$ 的 Borel 可测函数）……………………………… (232)

9—6　在一定的优良性准则下，用$S_{n+1}^2\triangleq\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 去估计 σ^2 较之用 $S_n^{*2}\triangleq\frac{1}{n-1}\sum_{i=1}^{n}\times(X_i-\bar{X})^2$去估计 σ^2 更优的例子………… (234)

9—7　无偏估计不是一致（相合）估计的例子…… (235)

9—8　无偏估计的方差低于 Rao-Cramér 不等式下界的例子……………………………………… (238)

9—9　UMVUE其方差达不到Rao-Cramér不等式下界的例子……………………………………… (240)

9—10　充分统计量不是有效估计量的例子………… (241)

(*)9—11　参数的UMVUE 不是参数的可容许估计的例子……………………………………………… (244)

9—12　某些条件不满足，但似然方程仍然存在一致

(相合)解并且满足渐近正态性的例子………(250)

(*)9--13 参数θ的一致（相合）渐近正态估计$\hat{\theta}$在θ点的渐近方差$\frac{v(\theta)}{n}$中的$v(\theta)$低于$C—R$下界的例子………(254)

9—14 若$T_n(X)$是θ的一致估计，又$|T_n-\theta|\leqslant A_n<\infty$，则$T_n$不是$\theta$的均方一致估计〔即$E(T_n-\theta)^2\nrightarrow 0,(n\to\infty)$〕的例子………(260)

9—15 极大似然估计不是充分统计量的例子………(262)

9—16 极大似然估计不是有效估计的例子………(266)

9—17 似然方程的解不是极大似然估计的例子……(267)

(*)9—18 极大似然估计不是一致估计的例子………(271)

第十章 与假设检验有关的反例………(277)

10—1 不是单参数指数族，但具有单调似然比的例子………(277)

(*)10—2 单边假设检验不存在一致最优势检验（记为UMP检验）的例子………(279)

(*)10—3 双边假设检验不存在 UMP 检验的例子……(282)

(*)10—4 检验函数 ϕ 对公共边界相似，但 ϕ 没有 Neyman 结构的例子………(287)

10—5 当正则条件不成立时，Wilks 定理中关于似然比极限分布的结论可以不成立的例子……(288)

10—6 对固定的样本大小而言，似然比检验可以不是无偏的例子………(291)

10—7 一个很坏的似然比检验的例子………(294)

第十一章 与线性模型有关的反例………(301)

11—1 在非正态的条件下，参数的最小方差线性无

偏估计不再是最小方差无偏估计的例子……（301）
第十二章　与抽样理论有关的反例……………………（306）
12—1　当检查人员可能犯错误时，全面检查不一定
比抽样检查更好的例子……………………（306）
主要参考文献……………………………………（312）

第一章　与随机事件、概率空间、古典概型有关的反例

1—1　基本事件但不是事件的例子

有关概念

1　随机试验的一个结果，称为一个基本事件．全体基本事件的集合称为样本空间（也称为基本事件空间）．

2　设Ω是样本空间，$\mathscr{F}$是由Ω的一些子集所组成的σ—域，则称$\mathscr{F}$为事件域，$\mathscr{F}$中的元素称为事件，Ω称为必然事件，ϕ称为不可能事件．

反例　在测量某一零件时，考虑其测量结果与真正长度的误差．从理论上讲，它可用一个实数x表示，故基本事件为$\{x\}$$(-\infty<x<\infty)$，整个实数轴为样本空间，即$\Omega=\{x|x\in R_1\}$．

若令$A_1=\{x|x\geqslant 0\}$，$A_2=\{x|x<0\}$，即A_1表示"出现非负误差"，A_2表示"出现负误差"．则可验证

$$\mathscr{F}\triangleq\{\phi,A_1,A_2,\Omega\}$$

是一个事件域．

由此可见，基本事件不一定是事件．

1—2 同一随机现象可以用不同的样本空间来描述的例子

有关概念 参看1—1

反例

例1 讨论检查n个产品这一随机现象.

若我们感兴趣的是产品中的次品个数k,那么它所有可能的基本事件有$n+1$个，即

$\omega_0=\{$没有次品$\}$，$\omega_1=\{$有一个次品$\}$，…，$\omega_n=\{$有n个次品$\}$，

所以它的样本空间

$$\Omega_1=\{\omega_0, \omega_1, \cdots, \omega_n\}.$$

若我们注意的是检查的记录结果,并记出现一个次品为1，出现一个正品为0，那么所有可能的基本事件有2^n个，它的样本空间为

$$\Omega_2=\{(0,0,\cdots,0), (1,0,\cdots,0), (0,1,0,\cdots,0), \cdots,(0,0,\cdots,0,1), (1,1,0,\cdots,0), (1,0,1,0,\cdots,0), \cdots,(0,\cdots,0,1,1), \cdots,(1,1,\cdots, 1)\}.$$

Ω_1中每一个基本事件$\omega_k=$“n个产品中有k个次品”($k=0, 1,2,\cdots,n$)是由Ω_2的某些基本事件组成的. 例如

$$\omega_1=\{(1,0,\cdots,0), (0,1,0,\cdots,0), \cdots,(0,\cdots,0,1)\}.$$

有的问题需要采用这里所指出的样本空间Ω_2. 比如,考察“第2个产品是次品”这一随机事件A. 此时用样本空间Ω_1就无

法讨论．但在Ω_2中

$$A=\{所有(*,1,*,\cdots,*)|其中*可为0或1\}\subset\Omega_2.$$

例2 袋中有a个黑球，b个白球，现在把球随机地一个个摸出来，求第k次摸出一个球是黑球的概率$(1\leqslant k\leqslant a+b)$。

观察把袋中$(a+b)$个球一个个摸出来这一随机试验．

如果我们假定“球有个性”，即认为同一种颜色的球之间仍有区别（例如设想把它们进行编号），若把摸出的球依次放在排列成一直线的$(a+b)$个位置上，每一种可能的排法就是一个样本点（即基本事件），总共有$(a+b)!$个样本点，它们就组成一个样本空间Ω_3。若用事件B表示第k次摸出一个球是黑球，则有利于B的样本点数有 $a\cdot(a+b-1)!$ 这是因为第k次摸得黑球有a种取法，而另外$(a+b-1)$次摸球相当于$(a+b-1)$个球进行全排列，有$(a+b-1)!$种构成法，故所求概率为

$$P_1=\frac{a\cdot(a+b-1)!}{(a+b)!}=\frac{a}{a+b}.$$

如果我们假定对同种颜色的球不加区别，即把a个黑球看作没有区别，把b个白球也看作没有区别，那么在$(a+b)$个位置上任取a个位置放上黑球（其他位置必然都放白球）就构成一个样本点，所有这些样本点（总数为C_{a+b}^{a}个）就组成样本空间Ω_4。在这种场合下，有利于事件A的样本点数共有C_{a+b-1}^{a-1}．这是因为第k次摸得黑球，这个位置必须放黑球，剩下的黑球可以放在$(a+b-1)$个位置中任取的$(a-1)$个位置上，所以共有C_{a+b-1}^{a-1}种放法，因此所求概率为

$$P_2=\frac{C_{a+b-1}^{a-1}}{C_{a+b}^{a}}=\frac{a}{a+b}.$$

两种不同的思考方法所得的结果相同．这两种解法的不同之处主要在于选取的样本空间Ω_3和Ω_4不同．Ω_4中的每一个基

本事件是Ω_3中的$a!b!$个样本点合并而成的。

本例表明，同一随机现象可以用不同的样本空间来描述，相应地同一随机事件的概率就可用不同的方法求出，但答案应该是相同的。

1—3 两事件互斥但不互逆的例子

有关概念

1 如果两事件A，B满足关系

$$A\cap B=\phi,$$

也就是说，如果A，B不能同时出现，则称A，B互斥（或称互不相容）。

2 如果两事件A，B满足关系

$$A\cup B=\Omega,\quad A\cap B=\phi.$$

也就是说，A，B中必出现一个，也只能出现一个，则称A、B互逆（或称A，B互为对立事件）。

显然，A，B互逆，必定A，B互斥，但反之不然。

反例 在0，1，2，3，…，9共十个数字中任意选取一个，有十个不同的结果："取得一个数字为0"，…，"取得一个数字为9"。每个结果都是基本事件。

令 事件A表示"取到一个数$\leqslant 3$"；

事件B表示"取到一个数$\geqslant 5$"，

则 $$A\cap B=\phi,$$

故A，B互斥。但是

$$\begin{aligned}A\cup B&=\{取得一个数\leqslant 3或\geqslant 5\}\\&=\{取得一个数为0,1,2,3,5,6,7,8,9\}\neq\Omega,\end{aligned}$$

所以A，B不互逆。

1—4 概率为0的事件并非是不可能事件的例子

有关概念

我们知道，不可能事件ϕ的概率为0，即$P(\phi)=0$，但其逆不真。

反例

例1 设A表示事件：向边长为a的正方形区域G任意投掷一点，此点落在区域g(g为该正方形的四条边)，则

$$P(A)=\frac{g\text{的面积}}{G\text{的面积}}=\frac{0}{a^2}=0,$$

但A却并非不可能事件。

例2 设ξ是连续型随机变量，取值充满R_1，其密度函数为$f(x)$。则对$\forall C\in R_1$，有

$$P(\xi=c)\leqslant P(c\leqslant\xi<c+h)=\int_c^{c+h}f(x)dx.$$

其中 $h>0$,

故 $$0\leqslant P(\xi=c)\leqslant\lim_{h\to 0}\int_c^{c+h}f(x)dx=0,$$

因此 $$P(\xi=c)=0,$$

即连续型随机变量取单个值的概率为0,然而$\{\xi=c\}$并非不可能事件。

1—5 概率为1的事件并非是必然事件的例子

有关概念

我们知道，必然事件Ω的概率$P(\Omega)=1$，但其逆不真。

反例

例1 以B表示事件：向边长为a的正方形G上任意投掷一点，此点落在区域g(g为正方形去掉一条对角线后所剩下的区域)，则

$$P(B)=\frac{g\text{的测度}}{G\text{的测度}}=\frac{a^2}{a^2}=1,$$

但B并非必然事件。

例2 设连续型随机变量ξ取值于$[a,b]$，c为$[a,b]$中任一点，则

$$P(\xi\in[a,b]-\{c\})=1,$$

但事件$\{\omega:\xi(\omega)\in[a,b]-\{c\}\}$并非必然事件。

(*) 1—6 集类E上的非负集函数μ具有有限可加性，连续性，但不具有可列可加性的例子

有关概念和命题

定义1 设μ是定义在集类$\mathbf{E}$上的广义实值集函数（即可以取有限或无限值的函数），如果对$\mathbf{E}$的任何不相交有限子集$\{E_1, E_2,\cdots,E_n\}$，若它的并集也在$\mathbf{E}$中，且有

$$\mu\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mu(E_i)$$

则称 μ 具有有限可加性；如果对**E**中之集的任何不相交序列$\{E_n\}$，若它的并集也在**E**中，且有

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i),$$

则称μ具有可列可加性。

定义2 设μ是定义在集类**E**上的广义实值集函数，如果对**E**中之集的每一个增序列$\{E_n\}$

有 $$\lim_n \mu(E_n) = \mu(E),$$

其中 $$\lim_n E_n = E,$$

则称μ在$E(E\in \mathbf{E})$是下连续的。类似地，如果对**E**中之集的每一个减序列$\{E_n\}$，有

$$\lim_n \mu(E_n) = \mu(E),$$

其中假定$\{E_n\}$中至少有一项E_m，使

$$|\mu(E_m)| < \infty,$$

且 $$\lim_n E_n = E,$$

则称μ在E是上连续的。

定义3 设$\mathscr{F}$是Ω的某些子集类，若满足

(i) $\Omega\in\mathscr{F}$；

(ii) 若$A\in\mathscr{F}$，则$\bar{A}\in\mathscr{F}$；

(iii) 若$A_n\in\mathscr{F}(n=1,2,3,\cdots)$，有

$$\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}.$$

则$\mathscr{F}$称为Ω的一个σ—域(又称事件域)。

定理 若P是σ—域$\mathscr{F}$上满足$P(\Omega)=1$的非负集合函数，则它具有可列可加性的充要条件是

(i) 它是有限可加的；

(ii) 它是下连续的。

推论 概率是下连续的。

反例

例1 设Ω是区间$[0,1]$中一切有理数的集，$\mathbf{E}$ 是形如$\{x|x\in\Omega,\ a\leqslant x<b\}$的一切半闭区间组成的类，其中 $0\leqslant a\leqslant b\leqslant 1$，$a$和$b$是有理数。在$\mathbf{E}$上定义集函数$\mu$如下：

$$\mu(\{x|a\leqslant x<b\})=b-a$$

则μ具有有限可加性，并且是下连续的。

事实上，设

$$E_i\in\mathbf{E},\qquad \bigcup_{i=1}^{n}E_i\triangleq E_0\in\mathbf{E}$$

诸E_i不相交，不妨设

$$E_i=[a_i,\ b_i)\quad (i=0,1,2,\cdots,n),$$

且 $$a_0=a_1\leqslant b_1=a_2\leqslant b_2=\cdots=a_n\leqslant b_n=b_0,$$

则 $$\mu(E_0)=\mu([a_0,b_0))=b_0-a_0$$

$$=b_n-a_1=\sum_{i=1}^{n}(b_i-a_i)$$

$$=\sum_{i=1}^{n}\mu(E_i),$$

即μ具有有限可加性。

再设

$$E_n=[a_n,b_n)\in\mathbf{E},\quad E_n\uparrow E_0=[a_0,b_0)\in\mathbf{E},$$

则 $$a_n\downarrow a_0,\quad b_n\uparrow b_0,$$

$$\mu(E_0)=b_0-a_0=\lim_n b_n-\lim_n a_n$$
$$=\lim_n(b_n-a_n)=\lim_n\mu(E_n),$$

故μ是下连续的。

但是μ不具有可列可加性。下面用反证法来证明之。

假若不然，即μ具有可列可加性。则对任

$$E\in\mathbf{E},\quad E_i\in\mathbf{E},$$

$$E\subset\bigcup_{i=1}^{\infty}E_i,$$

应有
$$\mu(E)\leqslant\sum_{i=1}^{\infty}\mu(E_i).$$

(见[15]，P.40(1))。因为[0,1]中全体有理数构成可列集，故可记为$\{x_1,x_2,\cdots\}$。取

$$E_i=\left[x_i,\quad x_i+\frac{1}{2^{i+1}}\right)\cap\Omega\qquad(i=1,2,3,\cdots),$$

则显然有 $\Omega=[0,1]\subset\bigcup_{i=1}^{\infty}E_i$。

但
$$1=\mu(\Omega)\leqslant\sum_{i=1}^{\infty}\mu(E_i)=\sum_{i=1}^{\infty}\frac{1}{2^{i+1}}=\frac{1}{2},$$

矛盾。

例2 设Ω是一切正整数集，并设$\mathbf{E}$是Ω中一切有限集E及其余集组成的类。对$E\in\mathbf{E}$，若E是有限集，令 $\mu(E)=0$； 若E是无限集，令$\mu(E)=\infty$。则集函数μ在 ϕ 是上连续的，且有限可加，但μ不具有可列可加性。

事实上，μ在$\mathbf{E}$中的有限可加性是显然的。

设 $E_n\downarrow\phi$，$E_n\in\mathbf{E}(n=1,2,3,\cdots)$。

且存在E_m，使$\mu(E_m)<\infty$，则E_m必为有限集，从而当$n>m$时E_m

都是有限集，

$$\mu(E_n)=0,$$

所以 $$\lim_n \mu(E_n)=0=\mu(\phi),$$

即μ在ϕ是上连续的。

但是，若设

$$E_n=\{n\}\quad(n=1,2,3,\cdots),$$

则每个 $\mu(E_n)=0$。

而 $$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\infty\neq 0=\sum_{n=1}^{\infty}\mu(E_n),$$

故μ不具有可列可加性。

进一步的讨论

本条中的定理表明：如果P是定义在Ω的σ—域$\mathscr{F}$上的非负集合函数，且满足 $P(\Omega)=1$， 则P具有可列可加性等价于它具有有限可加性并且是下连续（或者上连续）的。此处，对于集合函数的定义域的结构尚可作进一步的放宽，而不必要求它是σ—域

先给出一个定义：

一个以集为元素的非空类$\mathbf{R}$，若$E\in\mathbf{R}$，$F\in\mathbf{R}$，则 $E\cup F\in\mathbf{R}$，$E-F\in\mathbf{R}$，则称$\mathbf{R}$为环。

下述结果可以在〔15〕，$P.40$定理6找到。

定理 设μ是定义在环$\mathbf{R}$上的非负有限实值集函数，并具有有限可加性。如果μ在$\mathbf{R}$中每一个集E是下连续的 （或在ϕ是上连续的），则μ具有可列可加性。

注意到上述反例1中的集类 $\mathbf{E}$ 不是环（因为它对差运算不封闭）；反例2中的集类$\mathbf{E}$虽然是一个环，但 μ 可以取无限值。

在这两种情况下，由μ的有限可加性和连续性，都推不出 μ 的可列可加性。

1—7 上极限事件和下极限事件不相等的例子

有关概念

设A_1，A_2，…，A_n，…是一列事件，记

$$\overline{\lim_{n\to\infty}} A_n = \bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n,$$

$$\underline{\lim_{n\to\infty}} A_n = \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n,$$

称$\overline{\lim\limits_{n\to\infty}} A_n$为事件序列$\{A_n\}$的上极限事件，它表示 $A_n(n=1,2,3,\cdots)$发生无穷多次；称$\underline{\lim\limits_{n\to\infty}} A_n$为事件序列 $\{A_n\}$ 的下极限事件，它表示$A_n(n=1,2,3,\cdots)$ 中至多只有有限多个不发生。显然

$$\underline{\lim_{n\to\infty}} A_n \subset \overline{\lim_{n\to\infty}} A_n,$$

特别地，当

$$\underline{\lim_{n\to\infty}} A_n = \overline{\lim_{n\to\infty}} A_n$$

时，记 $\lim\limits_{n\to\infty} A_n = \underline{\lim\limits_{n\to\infty}} A_n = \overline{\lim\limits_{n\to\infty}} A_n$，

并称它们为事件序列$\{A_n\}$的极限事件。

下面提供一个$\underline{\lim\limits_{n\to\infty}} A_n \neq \overline{\lim\limits_{n\to\infty}} A_n$，从而$\lim\limits_{n\to\infty} A_n$ 不存在的例子。

反例 设

$$A_n = \begin{cases} A, & \text{当 } n \text{ 为偶数时}, \\ B, & \text{当 } n \text{ 为奇数时}. \end{cases}$$

且 $A \neq B$,

则 $\underline{\lim}_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n = A \cap B$,

$\overline{\lim}_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n = A \cup B$.

1—8 非古典型随机试验的例子

有关概念

如果随机试验E具有下列两性质：

(1) 只有有穷多个不同的基本事件ω_1, ω_2, …, ω_n;

(2) 所有基本事件都是等可能的。

则称E为古典型随机试验。

反例

例1 （基本事件个数有穷，但非等可能）

任意地选取一本英文书．翻开任意一页，查阅某行的某字符，它可能是26个字母中的某一个，也可能是空格或标点符号（为方便计，通称为"空格"）。于是，任意查看某一字符，基本事件总数是有限的，共27个．但是，它们并不是等可能的．例如，字母e出现的可能性远比字母z出现的可能性来得大。

例2 （基本事件个数无穷，但等可能）

在一个面积为S的海域里有表面积为 A 的大陆架贮藏着石油．假定在这海域里任意选定一点钻探，可以认为该海域中各点被选中的可能性是一样的，但所有可能结果总数是无限的。此时，等可能性是通过下列方式来赋予意义的：落在某区域$G \subset S$的概率与区域G的面积成正比，而与其位置及形状无关。

例3 （基本事件个数无穷，且不等可能）

从装有m个白球和 n 个黑球的罈子中将球一个一个地取出来（但每次都放回），一直取到白球为止。

这个随机试验的所有基本事件有：

$\omega_1=$（白），$\omega_2=$（黑，白），$\omega_3=$（黑，黑，白），…，

$\omega_n=$（黑，…，黑，白），…

↑

第 n 个

基本事件空间

$$\Omega=\{\omega_1,\ \omega_2,\ \cdots,\ \omega_n,\ \cdots\}.$$

用事件A_k表示"第k次取到黑球"

B_k表示"第k次取到白球"

则 $\bar{A}_k=B_k$，

$$P(A_k)=\frac{n}{m+n},\quad P(B_k)=\frac{m}{m+n}.$$

由于每次抽取后都放回，所以A_1，A_2，…，A_{k-1}，$\bar{A}_k$相互独立，因此

$$\begin{aligned}P(\omega_k)&=P(A_1,\ A_2,\ \cdots,\ A_{k-1},\ \bar{A}_k)\\&=P(A_1)\cdot P(A_2)\cdots P(A_{k-1})\cdot P(B_k)\\&=\left(\frac{n}{m+n}\right)^{k-1}\cdot\left(\frac{m}{m+n}\right)\qquad(k=1,2,3,\cdots)\end{aligned}$$

显然 $P(\omega_1)>P(\omega_2)>P(\omega_3)>\cdots>P(\omega_k)>\cdots$

可见诸基本事件ω_1，ω_2，…，ω_n，…不仅总数为无限，而且不是等可能的。

第二章　与独立性、条件概率有关的反例

2—1　两事件不独立的例子

有关概念

定义　对事件A和B，若

$$P(AB)=P(A)\cdot P(B),$$

则称它们是相互独立的，否则就称A，B不独立。

推论　若$P(A)>0$，则A，B独立的充分必要条件是

$$P(B|A)=P(B)。$$

反例

例1　袋中装有a个黑球和b个白球，采用不放回的摸球方式。

以　事件A表示“第一次摸得黑球”；

　　事件B表示“第二次摸得黑球”

此时　$P(A)=\dfrac{a}{a+b}$,

$$P(AB)=P(A)P(B|A)=\frac{a}{a+b}\cdot\frac{a-1}{a+b-1},$$

$$P(\bar{A}B)=P(\bar{A})\cdot P(B|\bar{A})$$

$$= \frac{b}{a+b} \cdot \frac{a}{a+b-1},$$

于是 $P(B) = P(AB) + P(\bar{A}B)$

$$= \frac{a}{a+b} \cdot \frac{a-1}{a+b-1} + \frac{b}{a+b} \cdot \frac{a}{a+b-1}$$

$$= \frac{a}{a+b}.$$

故有 $P(AB) \neq P(A) \cdot P(B)$,

即事件A和B不独立。

这是因为在第一次摸到黑球后，由于采用不放回摸法，从而使袋中球的组成成份发生了变化（黑球减少了），因而当然要影响（减少）第二次摸到黑球的概率。

例2 考虑有两个小孩的家庭．假定任一小孩为男孩或女孩是等可能的．

用事件A表示“任取一家庭至多有一个男孩”；事件B表示“任取一家庭有男孩，也有女孩”，则

$$P(A) = 3/4, \quad P(B) = 2/4,$$

$$P(AB) = P(B) = 2/4,$$

所以 $P(AB) \neq P(A) \cdot P(B)$.

于是事件A、B不独立。

进一步的讨论

例如考虑有三个孩子的家庭，其他假设条件以及事件A、B的含义均不变，则有

$$P(A) = 4/8, \quad P(B) = 6/8,$$

$$P(AB) = 3/8,$$

得 $P(AB) = P(A)P(B)$.

此时A、B是独立的。

本款例1中，如果改为有放回的摸球，则事件A，B也是独立的。对这一结论，读者很容易验证。

2—2 两事件不独立，但不一定互斥的例子

有关概念

1 两事件独立与不独立概念，见例2-1。

2 两事件互斥概念，见例1-3。

反例

考虑有两个小孩的家庭。假定任一小孩为男孩或女孩是等可能的。

以 事件A表示“任取一家庭至多有一男孩”；事件B表示“任取一家庭既有男孩，又有女孩”；

事件C表示“任取一家庭，两个孩子全是男孩或全是女孩”。

则 $P(A)=3/4$，$P(B)=2/4$，$P(C)=2/4$，

$P(AB)=2/4\neq P(A)P(B)$，

$P(BC)=0\neq P(B)\cdot P(C)$。

于是，A和B，B和C都不独立。但是

$A\cap B\neq\phi$，A、B不互斥，

$B\cap C=\phi$，B，C互斥。

此例表明，当两个事件不独立时，它们可以互斥，也可以不互斥

进一步的讨论

两个事件A、B相互独立，其实质是一个事件B出现的概率

与另一个事件A是否出现没有关系，但这并不意味着事件A、B本身完全无关。而若两个事件A、B互斥，则意味着A的出现必然导致B的不出现，因而此时B出现的概率与另一事件A是否出现密切相关。

它们有下面的性质：

若A、B两事件独立，且$P(A)>0$，$P(B)>0$则A、B必定不互斥。

这就是说，A、B独立与A、B互斥在$P(A)>0$,$P(B)>0$的场合下不能同时成立。

这个性质的证明是很简单的，留给读者。

2—3 事件两两独立，但并非相互独立的例子

有关概念

定义 对于三个事件A、B、C，若下列四个等式同时成立，则称它们相互独立：

$$P(AB)=P(A)P(B),$$
$$P(BC)=P(B)P(C),$$
$$P(CA)=P(C)P(A),$$
$$P(ABC)=P(A)P(B)P(C).$$

若前三个式子成立，则称A、B、C两两独立。显然相互独立必定两两独立。在以前很长一段时间里，人们认为三个事件只要两两独立（即前三个关系式成立），则这三个事件就是相互独立的（也即第四个关系式也一定成立）。其实不然，请看。

反例

例1 设$\Omega=(\omega_1, \omega_2, \omega_3, \omega_4)$

是等可能的．即

$$P(\omega_i)=\frac{1}{4} \qquad (i=1,\ 2,\ 3,\ 4).$$

令 $\quad A=(\omega_1,\ \omega_2),\ B=(\omega_1,\ \omega_3),\ C=(\omega_1,\ \omega_4),$

则 $\quad P(A)=P(B)=P(C)=\frac{1}{2}.$

由于 $\quad AB=AC=BC=(\omega_1),$

所以 $\quad P(AB)=P(AC)=P(BC)=\frac{1}{4}=\frac{1}{2}\cdot\frac{1}{2}$

$$=P(A)\cdot P(B)=P(B)\cdot P(C)=P(C)\cdot P(A),$$

因而三事件A、B、C是两两独立的．然而它们却不相互独立．这是因为

$$P(ABC)=P(\omega_1)=\frac{1}{4}\neq\frac{1}{8}$$
$$=P(A)\cdot P(B)\cdot P(C).$$

例2 有m^2块外形相同的木板，其中$(m-1)$块上写A，$(m-1)$块上写B，$(m-1)$块上写C，有一块上写ABC，其余的不写字．

今从m^2块中随机抽取一块．设抽到一块上写有A字的事件为A；写有B字的事件为B，写有C字的事件为C．则

$$P(A)=\frac{(m-1)+1}{m^2}=\frac{1}{m},$$

$$P(B)=\frac{(m-1)+1}{m^2}=\frac{1}{m},$$

$$P(C)=\frac{(m-1)+1}{m^2}=\frac{1}{m},$$

$$P(AB)=\frac{1}{m^2}, \qquad P(BC)=\frac{1}{m^2},P(CA)=\frac{1}{m^2},$$

所以 $P(AB)=P(A)\cdot P(B)$,

$$P(BC)=P(B)\cdot P(C),$$

$$P(CA)=P(C)\cdot P(A).$$

故A、B、C三事件两两独立。但是

$$P(ABC)=\frac{1}{m^2},$$

而 $$P(A)\cdot P(B)\cdot P(C)=\frac{1}{m^3},$$

所以 $P(ABC)\neq P(A)P(B)P(C)$。

这表明A、B、C三事件不相互独立。

进一步的讨论

事实上，存在n个事件A_1，A_2，…，A_n，使其中任何$n-1$个事件是相互独立的，但是这n个事件并不是相互独立的。（见《科学通报》1980年数学、物理学、化学专辑P90—92）

2—4 $P(ABC)=P(A)P(B)P(C)$，但A，B，C三事件并非两两独立的例子

有关概念 见例2-3

反例

例1 有一均匀正八面体，其第一、二、三、四面染上红色；第一、二、三、五面染上白色，第一、六、七、八面染上黑色。

现令

事件A = {抛一次正八面体，朝下的一面出现红色}，

$B=\{$抛一次正八面体，朝下的一面出现白色$\}$，

$C=\{$抛一次正八面体，朝下的一面出现黑色$\}$，

则 $$P(A)=P(B)=P(C)=\frac{4}{8}=\frac{1}{2},$$

$$P(ABC)=\frac{1}{8}=P(A)P(B)P(C).$$

但是 $$P(AB)=\frac{3}{8}\neq\frac{1}{4}=P(A)\cdot P(B),$$

可见A、B、C不两两独立。

例2 设

$$\Omega=(\omega_1,\ \omega_2,\ \omega_3,\ \omega_4),$$

令 $$P(\omega_1)=\frac{\sqrt{2}}{2}-\frac{1}{4},\quad P(\omega_2)=\frac{1}{4},$$

$$P(\omega_3)=\frac{3}{4}-\frac{\sqrt{2}}{2},\quad P(\omega_4)=\frac{1}{4}.$$

再设 $A=(\omega_1,\ \omega_3)$，$B=(\omega_2,\ \omega_3)$，$C=(\omega_3,\ \omega_4)$，

则 $$P(ABC)=P(\omega_3)=\frac{3}{4}-\frac{\sqrt{2}}{2},$$

$$\begin{aligned}P(A)&=P(\omega_1)+P(\omega_3)\\&=\left(\frac{\sqrt{2}}{2}-\frac{1}{4}\right)+\left(\frac{3}{4}-\frac{\sqrt{2}}{2}\right)\\&=\frac{1}{2},\end{aligned}$$

$$\begin{aligned}P(B)&=P(\omega_2)+P(\omega_3)\\&=1-\frac{\sqrt{2}}{2},\end{aligned}$$

$$\begin{aligned}P(C)&=P(\omega_3)+P(\omega_4)\\&=1-\frac{\sqrt{2}}{2},\end{aligned}$$

$$P(A)P(B)P(C)=\frac{1}{2}\left(1-\frac{\sqrt{2}}{2}\right)\left(1-\frac{\sqrt{2}}{2}\right)$$

$$=\frac{3}{4}-\frac{\sqrt{2}}{2},$$

故 $$P(ABC)=P(A)P(B)P(C).$$

但是 $$P(AB)=P(\omega_3)=\frac{3}{4}-\frac{\sqrt{2}}{2}$$

$$\neq P(A)P(B)=\frac{1}{2}\left(1-\frac{\sqrt{2}}{2}\right).$$

进一步的讨论

(1) 在定义n个事件A_1, A_2, …, A_n的独立性时，如果对任意$S(1\leqslant S\leqslant n)$, 任意$1\leqslant i_1<i_2<\cdots<i_s\leqslant n$, 有

$$P(A_{i_1}\cdot A_{i_2}\cdots A_{i_s})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_s})^{(*)}$$

则称事件A_1, A_2, …, A_n相互独立。

若对任意$1\leqslant i_1<i_2\leqslant n$, 有

$$P(A_{i_1}A_{i_2})=P(A_{i_1})P(A_{i_2}),$$

则称事件A_1, A_2, …, A_n 两两独立。

由此定义可知事件A_1, A_2, …, A_n相互独立一定两两独立。但从例2-3和例2-4可知A、B、C三事件两两独立不一定相互独立；而且由

$$P(ABC)=P(A)P(B)P(C)$$

也推不出A、B、C两两独立。只有当(*)中各式都成立时才能称A_1, A_2, …, A_n相互独立。(*)中共有

$$C_n^2+C_n^3+\cdots+C_n^n=2^n-n-1$$

个关系式。

(2) 此外若A与C独立，B与C独立，也未必有$A\cup B$, $A\cap B$,

$A-B$与C独立。

例如令

$$\Omega=\{1,\ 2,\ 3,\ 4\},\ \mathrm{P}(1)=\mathrm{P}(2)=\mathrm{P}(3)=\mathrm{P}(4)=\frac{1}{4}$$

$$A=\{1,\ 2\},\ B=\{1,\ 3\},\ C=\{1,\ 4\},$$

则不难验证A与C独立，B与C独立，但$A\cup B$，$A\cap B$，$A-B$均与C不独立。

2—5　说明$P(A)$，$P(A|B)$，$P(AB)$三者不同含义的例子

有关概念

1　设$(\Omega,\ \mathscr{F},\ P)$为一概率空间，$A\in\mathscr{F}$，$B\in\mathscr{F}$，且$P(B)>0$。在“事件$B$已经发生”条件下，事件$A$发生的条件概率$P(A|B)$定义为

$$P(A|B)=\frac{P(AB)}{P(B)}.$$

2　可以用下述观点来考虑条件概率：

概率空间已从$(\Omega,\ \mathscr{F},\ \mathrm{P})$变为$(\Omega_B,\ \mathscr{F}_B,\ \mathrm{P}_B)$，其中

$$\Omega_B=\Omega\cap B,\ \mathscr{F}_B=\{E\cap B|E\in\mathscr{F}\},$$

$$P_B(A)\stackrel{\wedge}{=}\frac{P(A)}{P(B)},\ \ \forall A\in\mathscr{F}_B.$$

例　设有5件产品，其中3件正品，2件次品。进行不放回抽样，连续抽两次。

以　事件B表示“头次抽到正品”；

　　事件A表示“第二次抽到正品”。

我们把连续两次抽取的结果看作为一个试验结果，并用a_1, a_2, a_3表示3件正品，用b_1，b_2表示2件次品。

于是，样本空间

$$\Omega=\left\{\begin{array}{l}(a_1,a_2),(a_2,a_1),(a_3,a_1),(b_1,a_1),(b_2,a_1)\\(a_1,a_3),(a_2,a_5),(a_3,a_2),(b_1,a_2),(b_2,a_2)\\(a_1,b_1),(a_2,b_1),(a_3,b_1),(b_1,a_3),(b_2,a_3)\\(a_1,b_2),(a_2,b_2),(a_3,b_2),(b_1,b_2),(b_2,b_1)\end{array}\right\}$$

此时

$$A=\left\{\begin{array}{lll}(a_1,a_2), & (a_2,a_1), & (a_3,a_1)\\(a_1,a_3), & (a_2,a_3), & (a_3,a_2)\\(b_1,a_1), & (b_1,a_2), & (b_1,a_3)\\(b_2,a_1), & (b_2,a_2), & (b_2,a_3)\end{array}\right\}$$

它包含Ω中十二个样本点，所以

$$P(A)=\frac{12}{20}=\frac{3}{5},$$

而 $$AB=\left\{\begin{array}{lll}(a_1, a_2), & (a_2, a_1), & (a_3, a_1)\\(a_1, a_3), & (a_2, a_3), & (a_3, a_2)\end{array}\right\}$$

所以 $$P(AB)=\frac{6}{20}=\frac{3}{10}.$$

在"事件B已经发生"条件下，样本空间已缩小为

$$\Omega_B=B\cap\Omega=\left\{\begin{array}{lll}(a_1,a_2), & (a_2,a_1), & (a_3,a_1)\\(a_1,a_3), & (a_2,a_3), & (a_3,a_2)\\(a_1,b_1), & (a_2,b_1), & (a_3,b_1)\\(a_1,b_2), & (a_2,b_2), & (a_3,b_2)\end{array}\right\}$$

在Ω_B的总共十二个样本点中，事件A只含有六个，所以

$$P(A|B)=\frac{6}{12}=\frac{P(A\cap B)}{P(B)}.$$

2—6 非独立试验概型(非Bernoulli概型)的例子

有关概念

设E为一随机试验,只有两个基本事件A和$\bar{A}$。令$P=P(A)$,$q=P(\bar{A})$,$0<p<1$,$0<q<1$,$p+q=1$。

将E独立地重复n次(或可列次),称为n重(或可列重)Bernoulli试验,并记为E^n(或E^∞)。这里的"重复",是指在每次试验中事件A,从而事件$\bar{A}$出现的概率都保持不变;这里的"独立"是指各次试验的结果互不影响。

反例 (Polya模型)

设罐中装有b个黑球,r个红球,任意取出一个,然后放回去,并再放入c个与取出的颜色相同的球。再向罐里取出一个球,如此继续下去,…。

用R_k表示"第k次取到红球"

用B_k表示"第k次取到黑球"

显然 $B_k=\bar{R}_k$

则 $P(R_1)=\frac{r}{b+r}$, $P(B_1)=\frac{b}{b+r}$.

下面用归纳法证明,对任何n,都有

$$P(R_n)=\frac{r}{b+r},\quad P(B_n)=\frac{b}{b+r}.$$

事实上,当$n=1$时,已有

$$P(R_1)=\frac{r}{b+r},\qquad P(B_1)=\frac{b}{b+r}.$$

假设 $P(B_k)=\frac{b}{b+r}$, $P(R_k)=\frac{r}{b+r}$,

往证 $P(R_{k+1})=\frac{r}{b+r}$， $P(B_{k+1})=\frac{b}{b+r}$，

由于 $P(R_{k+1})=P(R_{k+1}\cap R_1)+P(R_{k+1}\cap B_1)$

$$=P(R_1)P(R_{k+1}|R_1)+P(B_1)P(R_{k+1}|B_1).$$

考虑$P(R_{k+1}|R_1)$，在第一次取到红球后，罐内有红球$r+c$个，黑球b个，此时再抽出红球的概率将为$\frac{r+c}{b+r+c}$(抽出黑球的概率将为$\frac{b}{b+r+c}$)。根据归纳假设，在此后第k次试验抽到红球的概率仍旧等于此概率，故

$$P(R_{k+1}|R_1)=\frac{r+c}{b+r+c}.$$

同理 $P(R_{k+1}|B_1)=\frac{r}{b+r+c}$，

因此 $P(R_{k+1})=\frac{r}{b+r}\cdot\frac{r+c}{b+r+c}+\frac{b}{b+r}\cdot\frac{r}{b+r+c}$

$$=\frac{r}{b+r},$$

$$P(B_{k+1})=1-P(R_{k+1})=\frac{b}{b+r}.$$

这就是说，在各次试验中取到红球的概率和取到黑球的概率保持不变。但是容易看到，各次试验之间是不独立的。所以本模型不是Bernoulli概型。

(*)2—7 $\limsup\limits_{n\to\infty} E|\xi_n|<\infty$，但$\xi_n \xrightarrow{L_1} \xi$不成立的例子

有关结论

我们知道由$\xi_n\xrightarrow{L_1}\xi$(即$E|\xi_n-\xi|\xrightarrow[n\to\infty]{}0$)可推出

$$\limsup_{n\to\infty} E|\xi_n|<\infty$$

（见〔1〕独立性这一章P403），但其逆不真，即由

$$\limsup_{n\to\infty} E|\xi_n|<\infty,$$

并不能保证

$$\xi_n \xrightarrow{L_1} \xi.$$

我们有如下的反例。

例1 设η_1，η_2，…，η_n…是独立同分布的随机变量序列，且

$$P(\eta_n=0)=P(\eta_n=2)=\frac{1}{2} \qquad n=1, 2, \cdots$$

令 $$\xi_n=\prod_{i=1}^{n}\eta_i, \quad n=1, 2, \cdots$$

则不难证明$\{\xi_n\}$是鞅（鞅的概念见〔1〕P385）。自然$\{\xi_n\}$亦是下鞅。又

∵ $$E\xi_n=1$$

$\left(\text{故} \quad \limsup_{n\to\infty} E|\xi_n|=1<\infty\right)$.

∴ 由下鞅收敛性定理（见〔1〕P401）知存在随机变量ξ使

$$\xi_n \xrightarrow{a,s} \xi,$$

显然 $$\xi \overset{a,s}{=\!=} 0,$$

因此 $$E|\xi_n-\xi|=E\xi_n=1 \not\longrightarrow 0,$$

即 $$\xi_n \xrightarrow{L_1} \xi$$

不成立。

例2 令

$$\xi_n=(-1)^n,$$

则显然 $\varlimsup_{n\to\infty} E|\xi_n| = 1 < \infty$。

而对任意的ξ

$$E|\xi_n - \xi| \nrightarrow 0,$$

即 $\xi_n \xrightarrow{L_1} \xi$

不成立。

进一步的讨论

注 $\xi_n \to \xi \ a.s.$,

且 $\varlimsup_{n\to\infty} E|\xi_n| < \infty$

也未必有

$$\xi_n \xrightarrow{L_1} \xi.$$

例如令$\Omega = [0, 1]$，$\mathscr{F}$为$[0, 1]$中Borel点集全体，P为Lebesgue测度，$\xi \equiv 1$，

$$\xi_n = \begin{cases} n, & 当 0 < \omega \leqslant \dfrac{1}{n}, \\ 1, & 当 \dfrac{1}{n} < \omega \leqslant 1. \end{cases}$$

则显然有

$$\xi_n \longrightarrow \xi \ a.s.,$$

$$\varlimsup_{n\to\infty} E|\xi_n| \leqslant 2 < \infty.$$

但 $E|\xi_n - \xi| \longrightarrow 1 \neq 0$,

即 $\xi_n \xrightarrow{L_1} \xi$

不成立。

2—8 并非任何一个鞅$\{\xi_n\}$均存在一个满足$E|\xi|<\infty$的随机变量ξ和概率空间$(\Omega$、$\mathscr{F}_1, P)$中$\mathscr{F}_1$的子σ代数$\mathscr{D}_n$使得$\xi_n=E\{\xi|\mathscr{D}_n\}$的例子

有关定理

首先我们有如下的正面命题：

定理 设$\{\mathscr{F}_n\}$是$\mathscr{F}$中的非降子σ代数序列，ξ是一个随机变量，满足

$$E|\xi|<\infty$$

则
$$\xi_n=E\{\xi|\mathscr{F}_n\} \qquad (a,\ s) \qquad n\geqslant 1$$

是一个鞅（见〔1〕P389命题6.2.2)。但其逆不真，即给定一个鞅$\{\xi_n\}$，不一定存在随机变量ξ（满足$E|\xi|<\infty$)和$\mathscr{F}$中的非降子σ代数$\mathscr{F}_n$使

$$\xi_n=E\{\xi|\mathscr{F}_n\} \qquad (a,\ s).$$

在列举反例之前，我们先注意如下事实：设X_1，X_2，…，X_n…是独立的随机变量序列，

满足 $\quad EX_n=1 \quad$（对一切n)。

令
$$\eta_n=\prod_{i=1}^{n} X_i \qquad n\geqslant 1,$$

则不难证明$\{\eta_n\}$是一个鞅(〔1〕P387例6·2·3)。现在我们可以举出反例了。

反例 设η_1，η_2，…，η_n…是相互独立相同分布的随机变量序列，满足

$$P(\eta_i=0)=P(\eta_i=2)=\frac{1}{2} \quad i=1,\ 2,\ \cdots$$

则 $E\eta_i = 1$,

令 $\xi_n = \prod_{i=1}^{n} \eta_i$.

则由刚才列举的事实知$\{\xi_n\}$是一个鞅，但不存在满足$E|\xi|<\infty$的随机变量ξ及$\mathscr{F}$中之非降子σ-代数序列$\{\mathscr{D}_n\}$，使

$$\xi_n = E\{\xi | \mathscr{D}_n\} \qquad (a, s).$$

我们用反证法：若存在满足$E|\xi|<\infty$的随机变量ξ和$\mathscr{F}$中的非降子σ-代数序列$\{\mathscr{D}_n\}$使得

$$\xi_n = E\{\xi | \mathscr{D}_n\} \qquad (a, s)$$

则ξ_n是$\mathscr{D}_n$可测函数。因此

$$A_n = \{\xi_n = 0\} \in \mathscr{D}_n.$$

于是，由条件概率之定义得

$$0 = \int_{A_n} \xi_n dP_{\mathscr{D}_n} = \int_{A_n} E(\xi | \mathscr{D}_n) dP_{\mathscr{D}_n})$$

$$= \int_{A_n} \xi dP = \int_{\Omega} \xi I_{A_n} dP = E(\xi I_{A_n}),$$

但 $|\xi I_{A_n}| \leqslant |\xi|$,

而 $E|\xi| < \infty$,

又 $I_{A_n} \xrightarrow{a,s} 1$,

∴ 由Lebesgue控制收敛定理得

$$0 = E\xi I_{A_n} \Longrightarrow E\xi = E\xi_n = 1,$$

矛盾.

*2—9 ξ与ζ独立，但$E(\xi|\zeta, \zeta) \neq E(\xi|\eta)$的例子

有关结论

我们知道，若ξ与η独立，则

$$E(\xi|\eta)=E\xi(a,\ s),$$

但ξ与ζ独立，并不能推出

$$E(\xi,\ |\eta,\ \zeta)=E(\xi|\eta),$$

请看下面的反例.

反例 设

$$\Omega=[0,\ 1],\ \mathscr{F}=\beta_{\Omega}$$

(Ω中子集所组成的Borelσ域) P为$\Omega=[0,\ 1]$上的 Lebesgue 测度. 令

$$\xi=x_{[0;\frac{1}{2}]},\ \eta=x_{[0;\frac{3}{4})},\ \xi=x_{[\frac{1}{4};\frac{3}{4}]}.$$

则ξ与ζ独立，然而

$$E(\xi|\eta)=\begin{cases}\dfrac{2}{3}, & \omega\in\left[0,\ \dfrac{3}{4}\right),\\ 0, & \text{其他}.\end{cases}$$

$$E(\xi|h,\zeta)=\begin{cases}0, & \omega\in\left[\dfrac{3}{4},\ 1\right],\\ \dfrac{1}{2}, & \omega\in\left[\dfrac{1}{4},\ \dfrac{3}{4}\right),\\ 1, & \omega\in\left[0,\ \dfrac{1}{4}\right).\end{cases}$$

$$\therefore\quad E(\xi|\eta,\ \zeta)\neq E(\xi,\ \eta).$$

第三章　与随机变量，分布函数有关的反例

3—1 $\xi(\omega)$是定义在概率空间$(\Omega, \mathscr{F}, P)$上的单值实函数但它不是$(\Omega, \mathscr{F}, P)$上的随机变量的例子

有关概念

设$\xi(\omega)$是定义在概率空间$(\Omega, \mathscr{F}, P)$上的单值实函数，如果对于直线上任一波雷尔点集B，有

$$\{\omega: \xi(\omega)\in B\}\in\mathscr{F},$$

则称$\xi(\omega)$为随机变量.

反例　设$\Omega=[0, 1]$，$\mathscr{F}$是$[0, 1]$上的勒贝格可测集构成的σ-代数，

$$P(A)=m(A) \qquad (\forall A\in\mathscr{F}),$$

其中　$m(A)$是勒贝格集A的测度.

再设E是$[0, 1]$上的L-不可测集（其存在性的证明见[15] §16）.

考察函数

$$\xi(\omega)=\begin{cases}1, & \omega\in E,\\ 0, & \omega\in\Omega-E=\bar{E}.\end{cases}$$

对$\forall x\in R_1$

$$\{\omega:\xi(\omega)<x\}=\begin{cases}\Omega, & x>1,\\ \bar{E}, & 0<x\leqslant 1,\\ \Phi, & x\leqslant 0.\end{cases}$$

可见，当$0<x\leqslant 1$时

$$\{\omega:\xi(\omega)<x\}=\bar{E}\overline{\in}\mathscr{F},$$

因此函数$\xi(\omega)$不是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量。

(*)3—2 随机变量的勒贝格可测函数不一定是随机变量的例子

有关概念和命题

设$y=g(x)$是R_1到R_1上的映射。(1) 若对任给R_1中的波雷尔点集B_1，均有

$$g^{-1}(B_1)\triangleq\{x:g(x)\in B_1\}\in\beta_1,$$

其中 β_1为R_1上波雷尔σ-域，则称$g(x)$是一元波雷尔可测函数。

(2) 若

$$g^{-1}(B_1)\triangleq\{x:g(x)\in B_1\}\in\mathscr{L}_1,$$

其中 $\mathscr{L}_1$为R_1上勒贝格 σ-域，则称$g(x)$是一元勒贝格可测函数。

我们知道，若ξ是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，而$g(x)$是一元波雷尔可测函数，则$g(\xi(\omega))$是$(\Omega, \mathscr{F}, P)$上的随机变量．但是，当$g(x)$是一元勒贝格可测函数，则$(\xi(\omega))$就不一定再是随机变量了。

反例 设 $\Omega=[0, 1]$,

$\mathscr{F}$为$[0, 1]$中勒贝格σ-域，

P为L-测度。

下面我们将给出一个勒贝格可测的实函数$g(y)$，以及一个连续而且递增（在严格意义下）的函数$\xi(x)$，其中$0\leqslant x\leqslant 1$，使得$g(\xi(x))$为勒贝格不可测函数的例子。

以下分几步论述：

（1）对每一个$x\in[0, 1]$，令

$$x=\sum_{i=1}^{\infty}\frac{\alpha_i}{3^i}=0.\ \alpha_1\alpha_2\alpha_3\cdots$$

其中　$\alpha_i=0,\ 1$或$2,\ i=1,\ 2,\ 3.$

如果$\boldsymbol{c}$是康脱(Cantor)集，则当$x\in\boldsymbol{c}$时，有$\alpha_i=0$或$2,\ i=1,2,\cdots$. 设$\boldsymbol{n}=\boldsymbol{n}(x)$是使$\alpha_n=1$的第一个指标；如果这样的$n$不存在，这就是说，如果$x\in\boldsymbol{c}$，则令$n(x)=\infty$。定义函数$\psi$如下：

$$\psi(x)=\sum_{1\leqslant i<n}\frac{\alpha_i}{2^{i+1}}+\frac{1}{2^n}.$$

（2）若$0\leqslant x\leqslant y\leqslant 1$，则

$$o=\psi(o)\leqslant\psi(x)\leqslant\psi(y)\leqslant\psi(1)=1.$$

事实上，因

$$o=\sum_{i=1}^{\infty}o/3^i=0.00\cdots$$

故　$n(o)=\infty,$

$\psi(o)=0.$

又　$$1=\sum_{i=1}^{\infty}2/3^i=0.222\cdots,$$

故　$n(1)=\infty,$

所以　$\psi(1)=1.$

且对任$x,\ y\in[0, 1]$，显然有

$$o=\psi(o)\leqslant\psi(x),\ \psi(y)\leqslant\psi(1)=1.$$

若$x=y$，由于x的三进位表示式是唯一的，所以有

$$\psi(x)=\psi(y).$$

若 $x=0.\alpha_1\alpha_2\cdots<y=0.\beta_1\beta_2\cdots$

则存在自然数j，使对$1\leqslant i<j$时，$\alpha_i=\beta_i$且$\alpha_j<\beta_j$，因此，此时若$n(y)<j$，则必有

$$n(x)=n(y)=j.$$

由ψ的定义有 $\varphi(x)=\psi(y)$；若$n(y)=j$，则有$\beta_j=1$，$\alpha_j=0$，故$\psi(x)<\psi(y)$。若$n(y)>j$，则$n(x)\geqslant j$，则有 $\beta_j=0$或2，又因为$\alpha_j<\beta_j$，故必有$\beta_j=2$，$\alpha_j\leqslant 1$。此时

$$\begin{aligned}\psi(y)-\psi(x)=&\Big[\sum_{1\leqslant i<n(y)}\frac{\beta_i}{2^{i+1}}+\frac{1}{2^{n(y)}}\Big]\\&-\Big[\sum_{1\leqslant i\leqslant n(x)}\frac{\alpha_i}{2^{j+1}}+\frac{1}{2^{n(x)}}\Big]\\=&\sum_{1\leqslant i<j}\frac{\beta_i}{2^{i+1}}+\frac{\beta_j}{2^{j+1}}+\sum_{j<i<n(y)}\frac{\beta_i}{2^{i+1}}\\&+\frac{1}{2^{n(y)}}-\Big[\sum_{1\leqslant i<j}\frac{\alpha_i}{2^{i+1}}+\sum_{j\leqslant i<n(x)}\frac{\alpha_i}{2^{i+1}}\\&-\frac{1}{2^{n(x)}}\Big]=\frac{1}{2^j}+\sum_{j<i<n(y)}\frac{\beta_i}{2^{i+1}}\\&+\frac{1}{2^{n(y)}}-\sum_{j\leqslant i<n(x)}\frac{\alpha_i}{2^{i+1}}-\frac{1}{2^{n(x)}}\\\geqslant&\frac{1}{2^j}-\sum_{j\leqslant i<n(y)}\frac{\alpha_i}{2^{i+1}}-\frac{1}{2^{n(x)}}.\end{aligned}$$

上式当$n(x)=j$时，右边为0；当$n(x)>j$时，因为$\alpha_j\neq 1$，又$\alpha_j<\beta_j$故必有$\alpha_j=0$，故

$$\begin{aligned}\psi(y)-\psi(x)\geqslant&\frac{1}{2^j}-\sum_{j<i<n(x)}\frac{\alpha_i}{2^{i+1}}-\sum_{i=n(x)}^{\infty}\frac{1}{2^{i+1}}\\\geqslant&\frac{1}{2^j}-\sum_{i=j+1}^{\infty}\frac{1}{2^{i+1}}=0.\end{aligned}$$

综上所述，只要$0\leqslant x\leqslant y\leqslant 1$，就有

$$0=\varphi(o)\leqslant\psi(x)\leqslant\psi(y)\leqslant\psi(1)=1.$$

(3) $\psi(x)$在[0，1]上连续.

为了证明$\psi(x)$在[0，1]上的连续性，只要注意如果这个函数在某一点$x_0\in[0,1]$间断，那么由于$\psi(x)$的单调性以及当$0\leqslant x\leqslant y\leqslant 1$时，$0\leqslant\psi(x)\leqslant\psi(y)\leqslant 1$，可知函数在$x_0$处的左、右极限中$\psi(x_0-0)$和$\psi(x_0+0)$都存在，并且总有$0\leqslant\psi(x_0-0)$和$\psi(x_0+0)\leqslant 1$. 因此在线段$0\leqslant y\leqslant 1$上的子开区间$(\psi(x_0-0),\psi(x_0))$和$(\psi(x_0),\psi(x_0+0))$中至少有一个不含函数$\psi(x)$的任何值。但是，这是不可能的，因为，一切二进位有理数是函数$\psi(x)$的值，而二进位有理数在线段$0\leqslant y\leqslant 1$上处处稠密。于是函数$\psi(x)$没有任何间断点；即是说，函数在[0，1]上一切点都连续。

(4) 由(2)(3)知ψ是[0，1]上非降连续函数，

所以 $$x=\frac{1}{2}[y+\psi(y)]$$

是严格增连续函数，且$x\in[0,1]$，故存在严格增的连续反函数$y=\xi(x)$，$\xi(x)$显然是概率空间$(\Omega,\mathscr{F},P)$上的随机变量。

(5) 集$\xi^{-1}(\mathbf{C})$是L-可测集，并且有正的测度.

因为$\mathbf{C}$是康脱集，$\Omega-\mathbf{C}$是由可列个构成区间组成。而由ψ的定义，在每一构成区间上$\psi=$常数。所以

$$\psi(\Omega-\mathbf{C})\triangleq\{\psi(x):x\in\Omega-\mathbf{C}\}$$

是可列集，又

$$\begin{aligned}\xi^{-1}(\Omega-\mathbf{C})&\triangleq\{x:\xi(x)\in\Omega-\mathbf{C}\}\\&=\left\{\frac{1}{2}(y+\psi(y)):\ y\in\Omega-\mathbf{C}\right\}.\end{aligned}$$

由于ψ在每一构成区间上取常数，所以

$$x=\frac{1}{2}(y+\psi(y))$$

是一线性变换．又由于x是严格单调，故$\xi^{-1}(\Omega-\boldsymbol{C})$是由互不相交的开区间组成．利用勒贝格测度的平移不变性，可知每一区间的测度是变换前测度的$\frac{1}{2}$，注意到

$$P(\boldsymbol{C})=0,\ P(\Omega-\boldsymbol{C})=1,$$

故有 $$P(\xi^{-1}(\Omega-\boldsymbol{C}))=\frac{1}{2},$$

所以 $$P(\xi^{-1}(\boldsymbol{C}))=\frac{1}{2}.$$

(6) 因为$\xi^{-1}(\boldsymbol{C})$是L–可测集，且测度为正，所以存在L–不可测集$\xi^{-1}(M)\subset\xi^{-1}(\boldsymbol{C})$（见[15]，§16定理5），注意到$P(\boldsymbol{C})=0$，因此，作为0的测度$\boldsymbol{C}$的子集$M$必定是$L$–可测集．

(7) 设M是(6)中所述的集，取g为M的示性函数，即

$$g(x)\triangleq I_M(x)=\begin{cases}1, & x\in M,\\ 0, & x\overline{\in} M.\end{cases}$$

因为M是L–可测集，所以g是L–可测函数，但是

$$g(\xi(x))=I_M(\xi(x))=\begin{cases}1, & \xi(x)\in M,\\ 0, & \xi(x)\overline{\in} M,\end{cases}$$

$$=\begin{cases}1, & x\in\xi^{-1}(M)\\ 0, & x\overline{\in}[\xi^{-1}(M)]\end{cases}=I_{\xi^{-1}(M)}(x).$$

由于$\xi^{-1}(M)$为L–不可测集，所以$g(\xi(x))$不是L–可测函数．因为我们所设的概率空间$(\Omega,\ \mathscr{F},P)$中的$\mathscr{F}$取为$[0,\ 1]$上的勒贝格σ–域，所以$g(\xi(x))$不可能是概率空间$(\Omega,\ \mathscr{F},\ P)$上的随机变量．

综上所述，我们在所设概率空间$(\Omega,\ \mathscr{F},P)$上构造了一个随机变量$\xi(x)$和一个L–可测函数g（g不是波雷尔可测函数），并

且证明了$g(\xi(x))$不是$(\Omega, \mathscr{F}, P)$上的随机变量。

3—3 既非离散型，又非连续型的分布函数的例子

有关概念和命题

1 定理 设$F(x)\triangleq P(\xi<x)$为随机变量ξ的分布函数，则

(1) $F(x)$为单调不减；

(2) $F(x)$为左连续；

(3) $F(-\infty)\triangleq\lim\limits_{x\to-\infty}F(x)=0$，

$F(+\infty)\triangleq\lim\limits_{x\to+\infty}F(x)=1$。

而且上述定理的逆命题亦真。即有

2 若实值函数 $F(x)(x\in R_1)$ 满足上述定理中的（1）—（3），则存在概率空间上的随机变量ξ以$F(x)$为其分布函数。

反例

例1 设随机变量ξ服从[0，2]上均匀分布，又设

$$g(x)=\begin{cases}x, & 0\leqslant x<1,\\ 1, & 1\leqslant x\leqslant 2.\end{cases}$$

则随机变量$\eta=g(\xi)$既不是离散型，又不是连续型的。

事实上，由题意知η以概率1在[0，1]上取值。故当$y\leqslant 0$时，$F_\eta(y)=0$；$y>1$时，$F_\eta(y)=1$；$0<y\leqslant 1$时，

$$F_\eta(y)=P(g(\xi)<y)=P(\xi<y)=y/2,$$

所以
$$F_\eta(y)=\begin{cases}0, & y\leqslant 0,\\ y/2, & 0<y\leqslant 1,\\ 1, & y>1.\end{cases}$$

由于η取单点{1}的概率

$$P(\eta=1)=F_\eta(1+0)-F_\eta(1)=\frac{1}{2}\neq 0,$$

故η不是连续型随机变量．同时它也不是离散型随机变量，因为它的分布函数不是阶梯函数（见图3—3）。

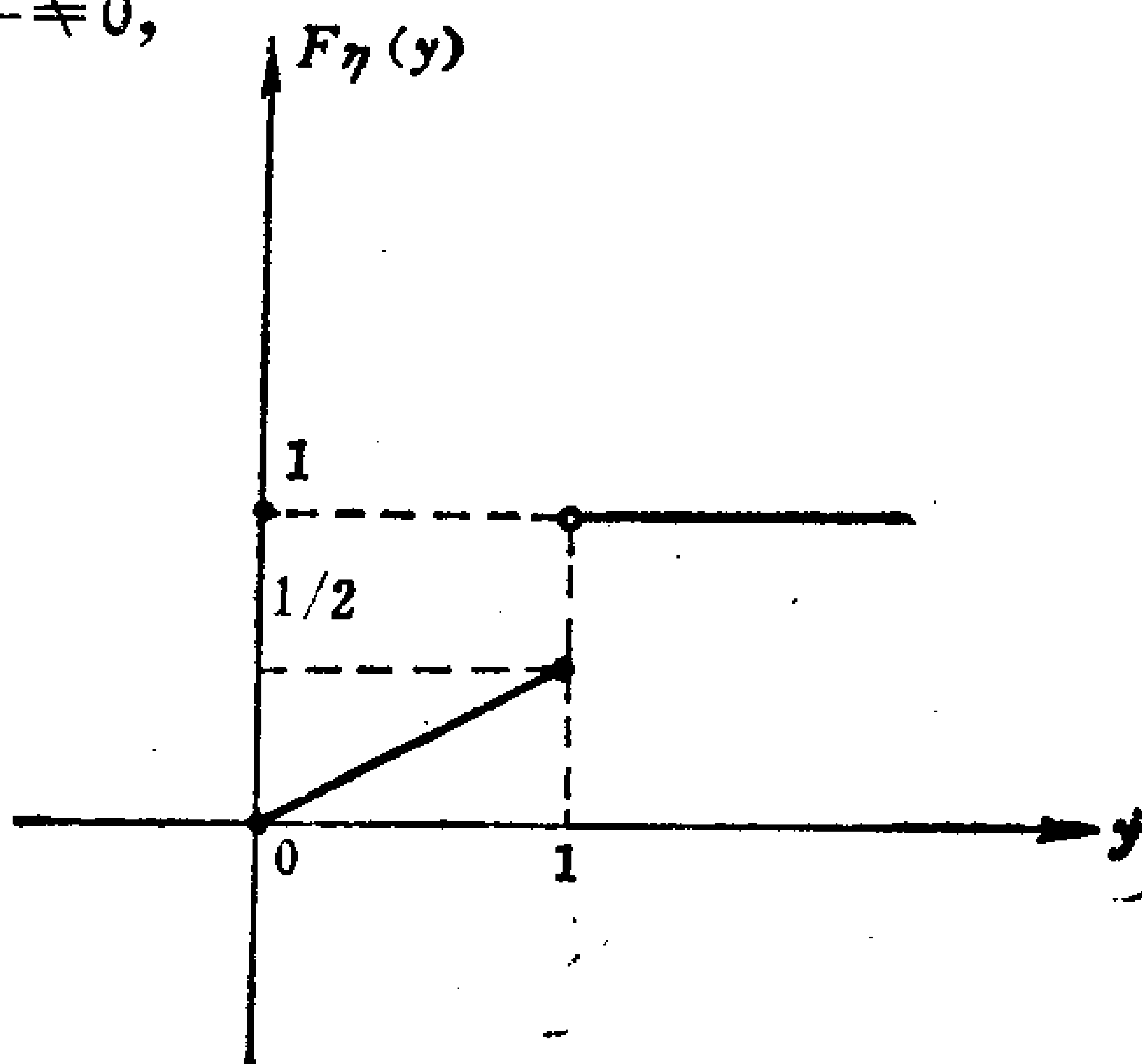


图3—3

(*)**例2**（奇异型分布函数）作康托（Cantor）函数$F(x)(x\in R_1)$，它的定义如下：

$$F(x)=\begin{cases}0, & \text{当}x<0,\\ 1, & x>1,\\ \frac{1}{2}, & x\in\left(\frac{1}{3},\ \frac{2}{3}\right),\\ \frac{1}{4}, & x\in\left(\frac{1}{9},\ \frac{2}{9}\right),\\ \frac{3}{4}, & x\in\left(\frac{7}{9},\ \frac{8}{9}\right),\\ \frac{1}{8}, & x\in\left(\frac{1}{27},\ \frac{2}{27}\right),\\ \frac{3}{8}, & x\in\left(\frac{7}{27},\ \frac{8}{27}\right),\\ \frac{5}{8}, & x\in\left(\frac{19}{27},\ \frac{20}{27}\right),\\ \frac{7}{8}, & x\in\left(\frac{25}{27},\ \frac{26}{27}\right).\end{cases}\qquad (*)$$

换句话说，把[0，1]分成三个等长的子区间，在中间第一个子区间$\left(\frac{1}{3},\ \frac{2}{3}\right)$上定义$F(x)=\frac{1}{2}$；把$\left(0,\ \frac{1}{3}\right)$，$\left(\frac{2}{3},\ 1\right)$各分成三

个等长的子区间，在两个中间子区间$\left(\frac{1}{9}, \frac{2}{9}\right)$（靠左的那个）和$\left(\frac{7}{9}, \frac{8}{9}\right)$（靠右的那个）上分别定义

$$F(x)=\frac{1}{2^2}$$

和
$$F(x)=\frac{3}{2^2}, \quad \cdots$$

依照这手续定义下去。一般地在第k秩（我们把上述过程中长为$\frac{1}{3^k}$的中间区间称为第k秩区间）的那些中间子区间上，从左往右算，在第k秩的第一个区间上定义$F(x)=\frac{1}{2^k}$，在第k秩的第二个区间上定义$F(x)=\frac{3}{2^k}$，在第k秩的第三个区间上定义$F(x)=\frac{5}{2^k}$，等等。在第k秩的最后一个区间上定义$F(x)=\frac{2^k-1}{2^k}$。因此，函数$F(x)$只在（$*$）右方诸区间的和的补集上未加定义，通常称此补集为康托集，记为C。函数$F(x)$在$\overline{C}=R_1-C$上单调，而且它的函数值所成之集在[0，1]上处处稠密（因为函数$F(x)$在$\overline{C}$上的值为介于0与1之间的一切二进位有理数）。

现在给函数在C上按照连续性的要求补充定义，即对$x\in C$，令

$$F(x)=\sup_{t<x,\, t\in\overline{C}} F(t),$$

于是$F(x)$在R_1上都有定义。

显然，$F(x)$ $(x\in R_1)$ 单调不减，连续，$F(-\infty)=0$，$F(+\infty)=1$，故它是一分布函数。因为它在$(-\infty, 0)(1, \infty)$

上各为常数，故 $F'(x)$在那里等于0；同理在$\left(\frac{1}{3},\ \frac{2}{3}\right)$，$\left(\frac{1}{9},\ \frac{2}{9}\right)$，$\left(\frac{7}{9},\ \frac{8}{9}\right)$，…等区间上 $F'(x)$也等于0。由于这些区间的总长度为1(见[15]，§15(5b))，故$F'(x)$几乎处处等于0(关于勒贝格测度而言)。

$F(x)$不可能是离散型分布函数,这是因为$F(x)$是连续函数，不是阶梯函数.又因为如果$F(x)$是连续型分布函数,则$\exists f(x)\geqslant 0(x\in R_1)$

使
$$F(x)=\int_{-\infty}^{x}f(t)dt,$$

于是几乎处处有
$$f(x)=F'(x)=0,$$

从而
$$\int_{-\infty}^{\infty}f(x)dx=\int_{-\infty}^{\infty}F'(x)dx=0$$

与
$$\int_{-\infty}^{\infty}f(x)dx=1$$

矛盾。故$F(x)$也不可能是连续型分布函数。

进一步的讨论

1 如果分布函数$F(x)$连续，而且它的导函数 $F'(x)$几乎处处为0(关于勒贝格测度而言)，则称此分布函数为奇异型的。

由本款反例2可知，补充定义后的康托函数 $F(x)$就是奇异型的分布函数。

任一奇异型分布函数$F(x)$，既不可能是离散型的，又不可能是连续型的。由此可知，没有一个分布函数能同时属于上述三种类型之二，故这三类分布函数是互相排斥的。

2 容易看出，如果a_1，a_2,a_3是三个非负常数,且a_1+a_2

$+a_3=1$，设$F_1(x)$，$F_2(x)$，$F_3(x)$分别是连续型，离散型，奇异型的分布函数，则$F(x)=a_1F_1(x)+a_2F_2(x)+a_3F_3(x)$（＊＊）也是一个分布函数。

事实上，由$F_i(x)(i=1, 2, 3)$的左连续性和单调不减性可知$F(x)$也有此二性质，又

$$\lim_{x\to-\infty}F(x)=a_1\lim_{x\to-\infty}F_1(x)+a_2\lim_{x\to-\infty}F_2(x)+a_3\lim_{x\to-\infty}F_3(x)=0,$$

$$\lim_{x\to+\infty}F(x)=a_1\lim_{x\to+\infty}F_1(x)+a_2\lim_{x\to+\infty}F_2(x)+a_3\lim_{x\to+\infty}F_3(x)=a_1+a_2+a_3=1.$$

而且，反面的结论也是正确的。即任一分布函数$F(x)$可唯一地展为（＊＊）的形式，其中$F_1(x)$，$F_2(x)$，$F_3(x)$分别为连续型，离散型，奇异型分布函数，而且$a_i\geqslant0(i=1, 2, 3)$，$a_1+a_2+a_3=1$。证明可参阅〔11〕，9.11或〔3〕中译本11.1。

3—4 设ξ是一个连续型随机变量，g是某个连续函数，$\eta=g(\xi)$不是连续型的随机变量的例子

有关概念

当随机变量ξ只能取有限个值$(x_1, x_2, \cdots, x_n)$或可列多个值$(x_1, x_2, \cdots)$时，称ξ为离散型随机变量。离散型随机变量的分布函数是一个阶梯函数。

当随机变量ξ的分布函数可表示为

$$F(x)=\int_{-\infty}^{x}f(t)dt,$$

其中$f(t)$为实数轴上的一个非负函数，则称ξ为连续型随机变

量.

连续型随机变量的分布函数一定是一个连续函数.

反例

设随机变量ξ在$[-1, 1]$上均匀分布,即它的分布密度函数为

$$f_{\xi}(x)=\begin{cases}\frac{1}{2}, & x\in[-1, 1],\\ 0, & x\bar{\in}[-1, 1].\end{cases}$$

令

$$\eta=\xi^{+}=\begin{cases}\xi, & \xi\geqslant 0,\\ 0, & \xi<0,\end{cases}$$

则η的分布函数

$$F_{\eta}(y)=P(\eta<y)$$

$$=p(\xi^{+}<y)=\begin{cases}0, & y\leqslant 0,\\ p(\xi<y), & y>0,\end{cases}$$

$$=\begin{cases}0, & y\leqslant 0,\\ \frac{1}{2}(1+y), & 0<y\leqslant 1,\\ 1, & y>1.\end{cases}$$

我们看到函数$F_{\eta}(y)$不是阶梯形跳跃,所以它是非离散型的,同时$F_{\eta}(y)$在$y=0$处有一个间断点,不连续,故不可能表为非负函数的变动上限的积分,因此$F_{\eta}(y)$又是非连续型的.

进一步的讨论

上例表明,当ξ是连续型随机变量时,为了保证$g(\xi)$同样也是一个连续型随机变量,还需要对g加某些条件.一个充分的条件由下列定理给出(见[2], $p70$).

定理 设ξ是一个连续型随机变量带有密度函数f.令函数

$y=g(x)$对所有x可微，并且或者对所有x，有$g'(x)>0$；或者对所有x，有$g'(x)<0$，则$\eta=g(\xi)$同样也是一个连续型随机变量，并且其密度函数由下式给出：

$$h(y)=\begin{cases} f[g^{-1}(y)]\left|\dfrac{d}{dy}g^{-1}(y)\right|, & \alpha<y<\beta, \\ 0, & \text{其余} \end{cases}$$

其中 $\alpha=\min\{g(-\infty),\ g(+\infty)\}$,

$\beta=\max\{g(-\infty),\ g(+\infty)\}$.

3—5 不同的随机变量(向量)，具有相同的分布函数的例子

有关概念

随机变量ξ的分布函数定义为

$$F(x)\hat{=}p(\omega;\ \xi(\omega)<x).$$

反例

例 1 设$\Omega=[0,\ 1]$，F是$[0,\ 1]$的Borel集类，p是通常的L—测度.

$$\xi_1(\omega)=-\xi_2(\omega)=\begin{cases} 1, & 0\leqslant\omega<\dfrac{1}{2} \\ -1, & \dfrac{1}{2}\leqslant\omega\leqslant 1. \end{cases}$$

则 $\xi_1(\omega)\neq\xi_2(\omega)$.

但是它们的分布函数却相同，都为

$$F(x)=\begin{cases} 0, & x\leqslant -1, \\ \dfrac{1}{2}, & -1<x\leqslant 1, \\ 1, & x>1. \end{cases}$$

例 2 （二维情况）不同的二维随机向量可以具有相同的（联合）分布函数．设Ω是正方形〔0，1；0，1〕，F是Ω的Borel集类，p是通常的二维L—测度．定义：

$$\xi_1(\omega)=\xi_2(\omega)=\begin{cases}-1, & \omega\text{位于}\Omega\text{的下半部分},\\ 1, & \omega\text{位于}\Omega\text{的上半部分}.\end{cases}$$

$$\eta_1(\omega)=\eta_2(\omega)=\begin{cases}-1, & \omega\text{位于}\Omega\text{的左半部分},\\ 1, & \omega\text{位于}\Omega\text{的右半部分}.\end{cases}$$

于是 $\{\omega|(\xi_1,\ \xi_2)\neq(\eta_1,\ \eta_2)\}$

$=\Omega$的左上及右下四分之一部分，

所以 $p\{\omega|(\xi_1,\ \xi_2)\neq(\eta_1,\ \eta_2)\}=\dfrac{1}{2}$.

但随机向量$(\xi_1,\ \xi_2)$和$(\eta_1,\ \eta_2)$的二维分布函数都是

$$F(x,\ y)=\begin{cases}0, & x,\ y\text{中有一个}\leqslant-1,\text{另一个任意},\\ \dfrac{1}{4}, & x,\ y\in-(1,\ 1],\\ \dfrac{1}{2}, & x,\ y\text{中一个在}(-1,\ 1],\text{另一个在}(1,\infty).\\ 1, & x,\ y\in(1,\ \infty).\end{cases}$$

3—6 连续型随机变量之密度函数未必是连续的例子

有关结论

我们知道，若随机变量ξ之密度函数是连续的，则ξ是连续型的随机变量．但其逆不真，我们有如下的

反例 设随机变量ξ的密度函数为

$$p(x)=\begin{cases}\frac{2}{3}x, & 当0\leqslant x<1,\\ 2(x-2)^2 & 当1\leqslant x\leqslant 2,\\ 0 & 其他.\end{cases}$$

则当$x\leqslant 0$时，ξ的分布函数

$$F(x)=\int_{-\infty}^{x}p(u)du=0$$

当$0<x\leqslant 1$时

$$F(x)=\int_{-\infty}^{x}p(u)du=\int_{-\infty}^{0}p(u)du+\int_{0}^{x}p(u)du$$

$$=0+\int_{0}^{x}\frac{2}{3}udu=\frac{x^2}{3}.$$

当$1<x\leqslant 2$时，

$$F(x)=\int_{-\infty}^{x}p(u)du=\int_{-\infty}^{0}p(u)du+\int_{0}^{1}p(u)du$$

$$+\int_{1}^{x}p(u)du$$

$$=0+\int_{0}^{1}\frac{2u}{3}du+\int_{1}^{x}2(u-2)^2du=1+\frac{2}{3}(x-2)^3.$$

当$x>2$时，$F(x)=1$.

此时密度函数与分布函数

$$F(x)=\begin{cases}0, & 当x\leqslant 0,\\ \frac{x^2}{3}, & 当0<x\leqslant 1,\\ 1+\frac{2}{3}(x-2)^3, & 当1<x\leqslant 2,\\ 1, & 当x>2.\end{cases}$$

的图象如图 3—6(1) 和3—6(2)。由此看到连续型随机变量ξ的密度函数未必是连续的。

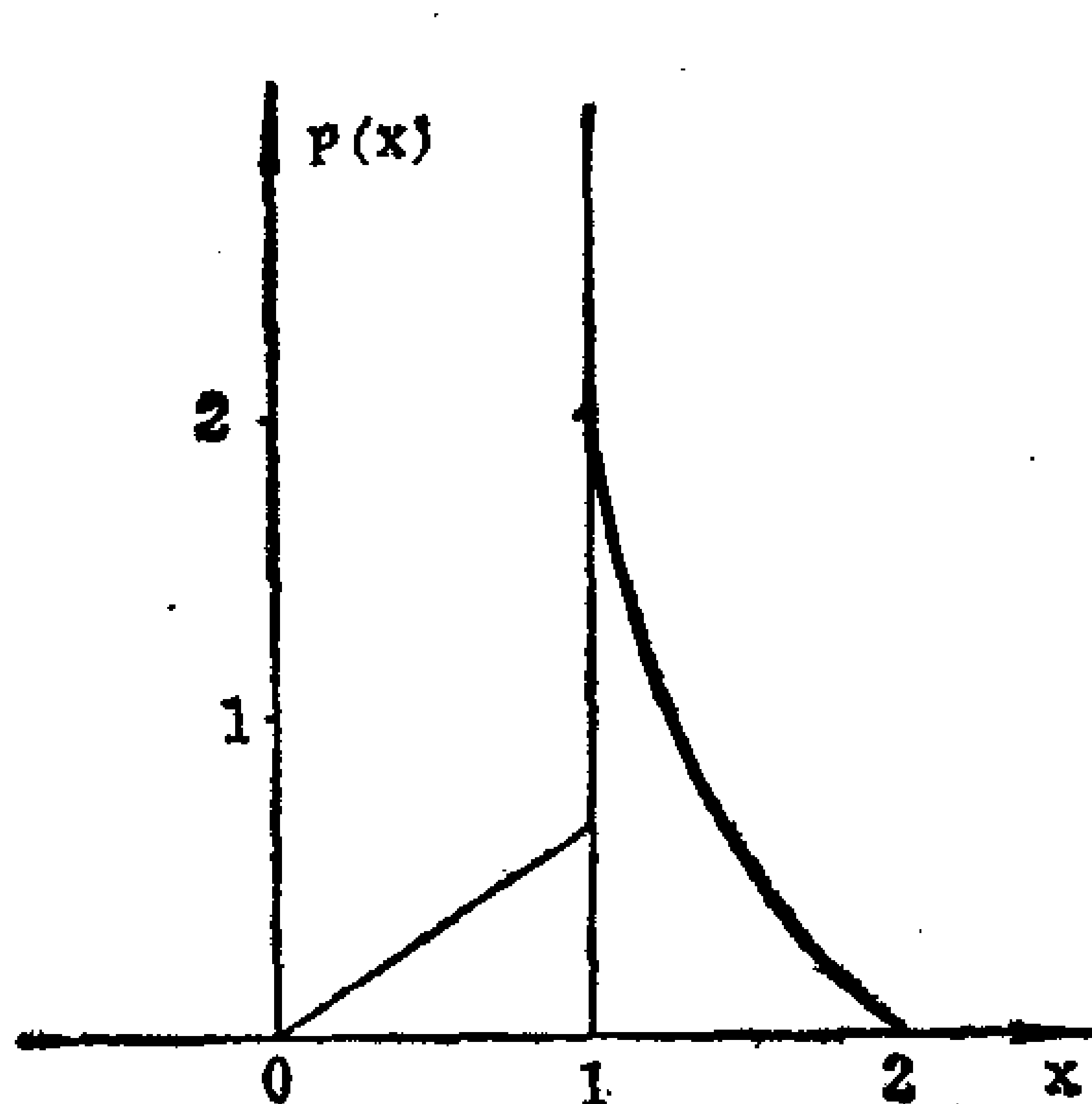


图3—6(1)

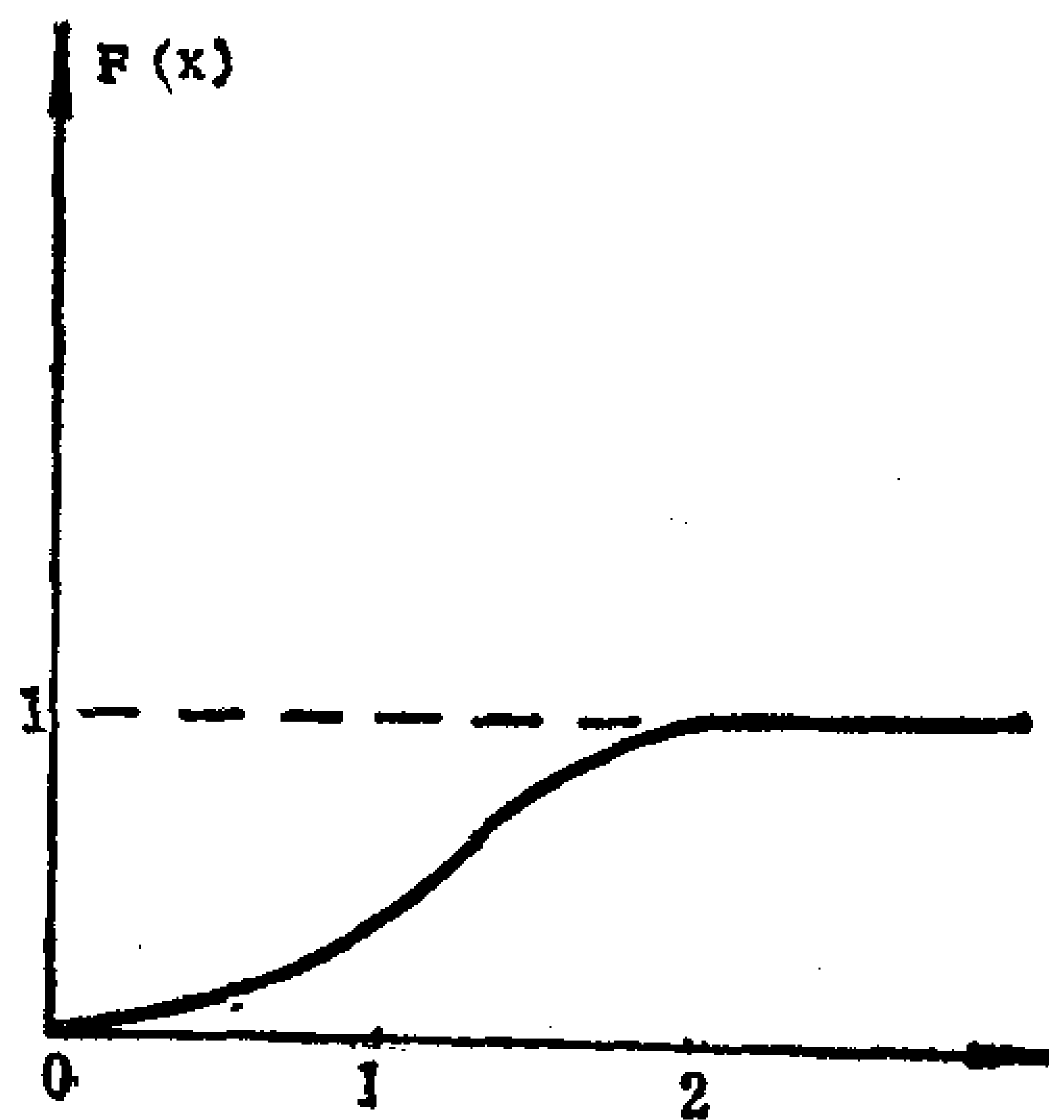


图3—6(2)

3—7 二元函数$F(x,y)$对每个变元非降，左连续，且$F(x,-\infty)=0$，$F(-\infty,y)=0$，$F(+\infty,+\infty)=1$，但仍不是分布函数的例子

有关概念

1 定理 设$F(x,\ y)$为随机向量$(\xi_1,\ \xi_2)$的联合分布函数，则

(1) $F(x,\ y)$分别对x，y单调非降；

(2) $F(x,\ y)$对每个变量左连续；

(3) $F(x,\ -\infty)=\lim\limits_{y\to-\infty}F(x,\ y)=0$,

$F(-\infty,\ y)=\lim\limits_{x\to-\infty}F(x,\ y)=0$,

$F(+\infty,\ +\infty)=\lim\limits_{\substack{x\to+\infty\\ y\to+\infty}}F(x,\ y)=1$。

(4) 对任意四个实数$a_1\leqslant b_1$，$a_2\leqslant b_2$，有

$p(a_1\leqslant\xi_1<b_1,\ a_2\leqslant\xi_2<b_2)$

$$=F(b_1, b_2)-F(a_1, b_2)-F(b_1, a_2)+F(a_1, a_2)\geqslant 0.$$

上述定理的逆命题也是成立的．即

2 若二元实值函数$F(x, y)(x, y\in R_1)$，满足上述定理中的(1)—(4)，则必存在随机变量ξ_1和ξ_2（在某个特定的概率空间上），使$F(x, y)$是(ξ_1, ξ_2)的联合分布函数．（证明参阅〔5〕，§2.6）

值得注意的是：在二维随机变量时与一维随机变量时不一样．刻划一个联合分布函数需要加条件(4)．这就是说，从条件(1)—(3)推不出(4)．

反例

例1 设

$$F(x, y)=\begin{cases}1, & 当x+y>-1,\\ 0, & 当x+y\leqslant -1.\end{cases}$$

显然，$F(x, y)$满足(1)—(3)，但不满足(4)．因为

$$F(1, 1)-F(1, -1)-F(-1, 1)+F(-1, -1)$$
$$=1-1-1+0=-1.$$

例2 定义二元函数

$$F(x, y)=\begin{cases}0, & x\leqslant 0或x+y\leqslant 1或y\leqslant 0,\\ 1, & 其他.\end{cases}$$

则$F(x, y)$满足上述条件(1)—(3)，然而$F(x, y)$不是分布函数．因为

$$F(1, 1)-F\left(1, \frac{1}{3}\right)-F\left(\frac{1}{3}, 1\right)+F\left(\frac{1}{3}, \frac{1}{3}\right)$$
$$=1-1-1-0=-1<0.$$

进一步的讨论

在本款反例2中，对$F(x, y)$我们不难找到无限多组不满

足(4)的点。因为从图3—7中可见，若点(x, y)位于第一象限中直线$x+y=1$的上方，则$F(x, y)=1$，否则，$F(x, y)=0$。因此我们只要任取一个边平行于坐标轴的矩形，其左下角的顶点在直线下方，其余三个顶点都位于第一象限中的直线上方，则都可以使(4)不满足。

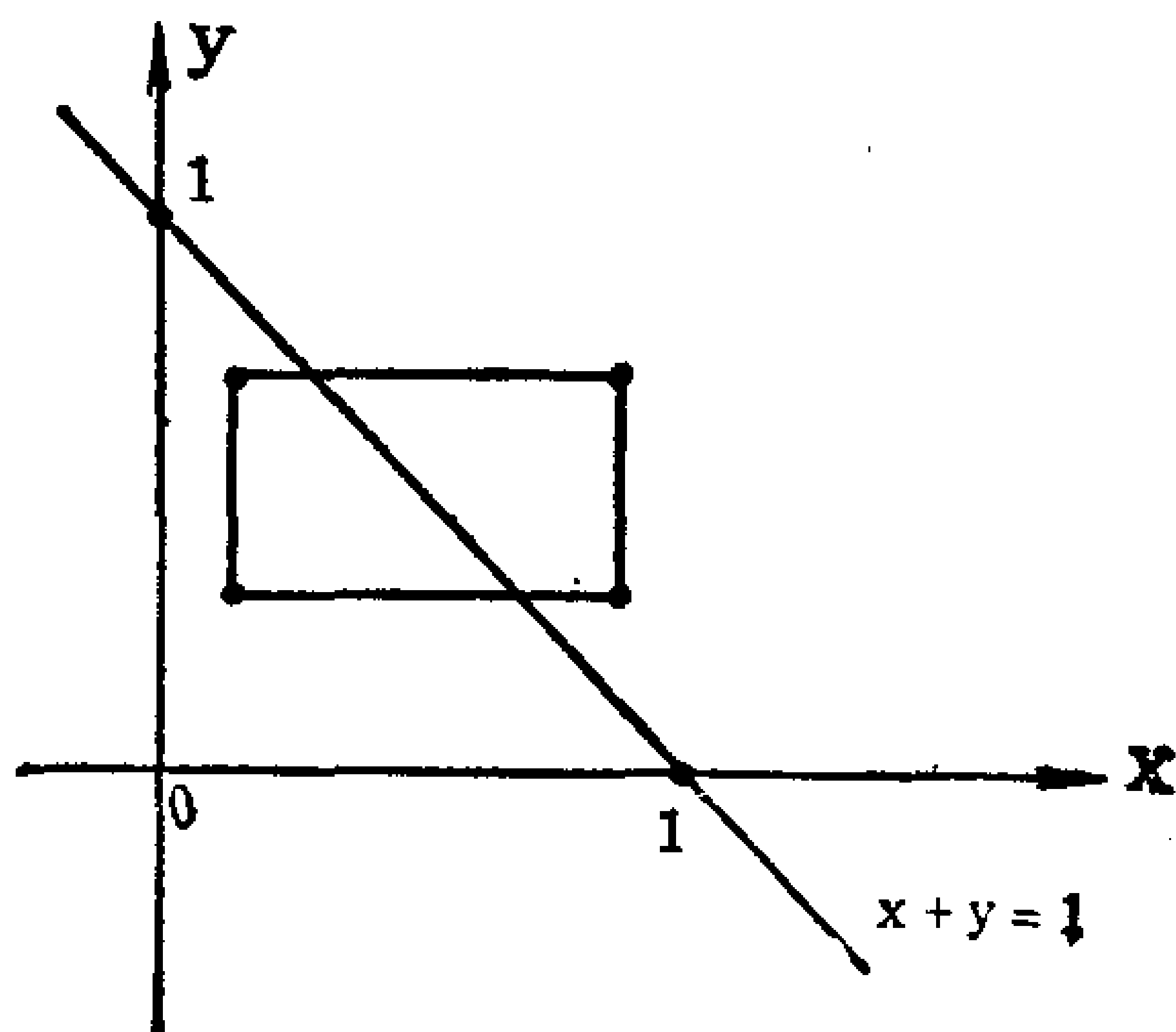


图3—7

3—8 边际分布是正态分布，但联合分布不是多元正态分布的例子

有关概念

函数

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\} \qquad (*)$$

(这里a，b，σ_1，σ_2，r为常数，$\sigma_1>0$，$\sigma_2>0$，$|r|<1$)称为二元正态分布密度函数。

(1) 二元正态分布(*)的边际分布仍为正态分布。

(2) 服从二元正态分布(*)的随机变量ξ_1, ξ_2独立的充要条件是$r=0$。这表明，对于二元正态分布，独立性和不相关性是等价的。

反例

例1 令

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (x \in R_1).$$

$$g(x) = \begin{cases} \cos x, & |x| < \pi, \\ 0, & |x| \geqslant \pi. \end{cases}$$

$$f(x, y) = \varphi(x)\varphi(y) + \frac{1}{2\pi} e^{-\pi^2} g(x) g(y) \quad (x \in R_1, \ y \in R_1),$$

则不难验证：

(1) $f(x, y)$是二元密度函数。即

$$f(x, y) = \begin{cases} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} + \frac{1}{2\pi} e^{-\pi^2} \cos x \cos y, & \left(\begin{matrix} |x| < \pi \\ |y| < \pi \end{matrix}\right), \\ \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} & (\text{其他}). \end{cases}$$

满足 $f(x, y) \geqslant 0$. $(x \in R_1, \ y \in R_1)$,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dx\,dy + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{1}{2\pi} e^{-\pi^2} \cos x \cos y\,dx\,dy$$

$$= 1 + \frac{1}{2\pi} e^{-\pi^2} \left[\int_{-\pi}^{\pi} \cos x\,dx \cdot \int_{-\pi}^{\pi} \cos y\,dy\right]$$

$$= 1.$$

（2）$f(x, y)$的边际分布都是正态分布：

$$f_{\xi_1}(x)=\int_{-\infty}^{\infty} f(x, y)dy$$

$$=\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dy+\int_{-\pi}^{\pi}\frac{1}{2\pi}e^{-\pi^2}\cos x\cos y dy$$

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

同理可得

$$f_{\xi_2}(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}.$$

（3）ξ_1，ξ_2不独立。这是因为

$$f(x, y)\neq f_{\xi_1}(x)\cdot f_{\xi_2}(y).$$

例如 $\quad f(0,0)=\frac{1}{2\pi}+\frac{1}{2\pi}e^{-\pi^2}$，

但 $\quad f_{\xi_1}(0)\cdot f_{\xi_2}(0)=\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{2\pi}}=\frac{1}{2\pi}\neq f(0, 0).$

（4）相关系数$r_{\xi_1\xi_2}=0$。因为，

$$E\xi_1=0,\ E\xi_2=0,\ D\xi_1=D\xi_2=1,$$

所以 $\quad \operatorname{cov}(\xi_1, \xi_2)=E(\xi_1, \xi_2)$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x, y)dxdy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy$$

$$+\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}e^{-\pi^2}xy\cos x\cos y dxdy$$

$$=\left(\int_{-\infty}x\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\right)\left(\int_{-\infty}^{\infty}y\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right)$$

$$+\frac{1}{2\pi}e^{-\pi^2}\left(\int_{-\pi}^{\pi}x\cos x\,dx\right)\left(\int_{-\pi}^{\pi}y\cos y\,dy\right)$$

$$=0,$$

故 $$r_{\xi_1\xi_2}=\frac{\operatorname{cov}(\xi_1,\ \xi_2)}{\sqrt{D\xi_1}\cdot\sqrt{D\xi_2}}=0.$$

(5) $f(x,\ y)$不是二元正态密度函数．因为如果$f(x,\ y)$是二元正态密度函数，则由$r_{\xi_1\xi_2}=0$要推出ξ_1，ξ_2独立，但此处ξ_1，ξ_2不独立，所以$(\xi_1,\ \xi_2)$的分布密度函数$f(x,\ y)$不可能是二元正态密度函数．

例 2 设随机变量$(\xi_1,\ \xi_2)$有如下的联合密度函数

$$f(x,y)=\frac{1}{2}\left\{\frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\exp\left[-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right]\right.$$
$$\left.+\frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\exp\left[-\frac{1}{2(1-\rho^2)}(x^2+2\rho xy+y^2)\right]\right\},$$

则不难验证它的两个边际密度函数都是正态密度函数；ξ_1和ξ_2的相关系数为0。事实上，

$$\xi_1\sim f_{\xi_1}(x)$$
$$=\frac{1}{2}\ \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}e^{-\frac{1}{2(1-\rho^2)}[(y-\rho x)^2+x^2(1-\rho^2)]}dy$$
$$+\frac{1}{2}\ \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}e^{-\frac{1}{2(1-\rho^2)}[(y+\rho x)^2+x^2(1-\rho^2)]}dy$$
$$=\frac{1}{2}\ \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}dy$$
$$+\frac{1}{2}\ \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}e^{-\frac{(y+\rho x)^2}{2(1-\rho^2)}}dy$$
$$=\frac{1}{2}\ \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\frac{1}{2}\ \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

类似可得

$$\xi_2 \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}},$$

而 $$E\,\xi_1\xi_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\left\{\frac{1}{2}\left[\frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}\right.\right.$$

$$\left.\left.+\frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2+2\rho xy+y^2)}\right]\right\}dxdy$$

$$=\frac{1}{4\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy\int_{-\infty}^{\infty} xye^{-\frac{1}{2(1-\rho^2)}(x-\rho y)^2}dx$$

$$+\frac{1}{4\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy\int_{-\infty}^{\infty} xye^{-\frac{1}{2(1-\rho^2)}(x+\rho y)^2}dx.$$

在上式右边的第一个积分中令

$$t=y,$$

$$z=\frac{1}{\sqrt{1-\rho^2}}(x-\rho y).$$

在上式右边的第二个积分中令

$$v=y,$$

$$u=\frac{1}{\sqrt{1-\rho^2}}(x+\rho y),$$

则有 $$E\xi_1\xi_2=\frac{1}{4\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sqrt{1-\rho^2}\,z+\rho t)\,t\,e^{-\frac{z^2}{2}-\frac{t^2}{2}}$$

$$\cdot\sqrt{1-\rho^2}\,dzdt$$

$$+\frac{1}{4\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sqrt{1-\rho^2}\,u-\rho v)\,ve^{-\frac{u^2}{2}-\frac{v^2}{2}}$$

$$\cdot\sqrt{1-\rho^2}\,dudv$$

$$=\frac{\rho}{4\pi}\int_{-\infty}^{\infty}t^2e^{-\frac{t^2}{2}}dt\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz+\frac{\sqrt{1-\rho^2}}{4\pi}\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt$$

$$\cdot\int_{-\infty}^{\infty} ze^{-\frac{z^2}{2}}dz-\frac{\rho}{4\pi}\int_{-\infty}^{\infty} v^2e^{-\frac{v^2}{2}}dv\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}du$$

$$+\frac{\sqrt{1-\rho^2}}{4\pi}\int_{-\infty}^{\infty} ue^{-\frac{u^2}{2}}du\int_{-\infty}^{\infty} ve^{-\frac{v^2}{2}}dv$$

$$=\frac{\rho}{2}+0-\frac{\rho}{2}+0=0.$$

$\therefore$ $$\text{cov}(\xi_1,\ \xi_2)=E\xi_1\xi_2-E\xi_1E\xi_2=0,$$

$$\rho=\frac{\text{cov}(\xi_1,\xi_2)}{\sqrt{D\xi_1}\sqrt{D\xi_2}}=0.$$

但ξ_1与ξ_2不独立，所以$f(x,\ y)$不可能是二元正态密度。

3—9　随机变量ξ，η相互独立，而且同分布，但不一定有$\xi=\eta(a,\ s)$成立的例子

有关概念

（a,s概念，又称“几乎处处”概念）

设$p(w)$是与自变数w有关的命题，如果使这个命题不成立的点w的全体只是一个零测度集，我们就称这个命题p“几乎处处”成立，也称a，s成立。

初学者常常会认为，随机变量ξ和η既然是同分布的，那么它们取的值应该一样，而且它们取同一个值的概率也是一样的，因而似乎当然有$\xi=\eta$了。其实此时不仅推不出它们相等，甚至连它们几乎处处相等都做不到。

反例

设随机变量ξ和η相互独立，而且它们的概率分布都是

$$\begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

此时因为ξ和η是独立同分布，所以ξ和η可能取的值都是一样的，都可能取-1和1两个值，而且它们取同一个值的概率确是一样的。但是由于ξ、η是随机变量，所以ξ取-1时，η不一定取-1，而可能取1；或者当ξ取1时，η可能取-1；当然也可能ξ和η同时取-1或同时取1。这正是随机变量与普通变量的本质区别。根据ξ和η的独立性假设，我们不难得$(\xi,\ \eta)$的联合分布：

ξ \ η	-1	1
-1	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$

由此可得

$$p(\xi \neq \eta) = p(\xi = -1\ \ \eta = 1) + p(\xi = 1,\ \ \eta = -1)$$

$$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

因此得不到$\xi = \eta(a,\ s)$。

3—10　ξ，η同分布但不独立时，$\zeta = \xi - \eta$ 不一定是对称随机变量的例子

有关概念

1　**定义1**　一个随机变量称为对称的，如果

$$p(\xi < x) = p(-\xi < x)$$

2　**定理**　若ξ、η是相互独立、相同分布的随机变量，则$\zeta = \xi - \eta$是一个对称的随机变量。

但是，当定理中ξ、η同分布的条件保留而相互独立的条件去掉时，$\zeta=\xi-\eta$不一定是一个对称的随机变量．我们有如下的

反例

例 1 设$(\xi,\ \eta)$的联合概率分布为

η \ ξ	1	2	3	$p(\eta=j)$
1	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{4}{12}$
2	$\frac{1}{12}$	0	$\frac{1}{12}$	$\frac{2}{12}$
3	$\frac{2}{12}$	0	$\frac{4}{12}$	$\frac{6}{12}$
$p(\xi=i)$	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{6}{12}$	1

$(i,\ j=1,\ 2,\ 3)$

显然，ξ、η是同分布的．但由于

$$p(\xi=1,\ \eta=1)=\frac{1}{12}\neq p(\xi=1)p(\eta=1)=\left(\frac{4}{12}\right)^2$$

故ξ、η不相互独立．

令$\zeta=\xi-\eta$，则ζ的概率分布为

$$\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \frac{1}{6} & \frac{1}{12} & \frac{5}{12} & \frac{3}{12} & \frac{1}{12} \end{pmatrix}$$

而$-\zeta=\eta-\xi$的概率分布为

$$\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \frac{1}{12} & \frac{3}{12} & \frac{5}{12} & \frac{1}{12} & \frac{2}{12} \end{pmatrix}$$

由于ζ和$-\zeta$没有相同的分布，故$\zeta=\xi-\eta$不是对称的随机变量．

但也可以找出这样的例子，ξ、η 同分布但不独立，然而

$\zeta=\xi-\eta$是对称的，请看下例。

例2 设

$$(\xi,\ \eta)\sim p(x,\ y)=\begin{cases}(1+a)-2a(x+y-2xy), & \\ \qquad 当0<x<1,\ 0<y<1, & \\ 0\quad ，其他。\end{cases}$$

其中$|a|\leqslant 1$，

则 $$\xi\sim p_1(x)=\begin{cases}1,\ 0<x<1, \\ 0,\ 其他.\end{cases}$$

同理 $$\eta\sim p_2(y)=\begin{cases}1,\ 0<y<1, \\ 0,\ 其他.\end{cases}$$

所以ξ、η同分布但不独立。

令$\zeta=\xi-\eta$，则

$$\zeta\sim p_3(z)=\begin{cases}\left(1+\dfrac{1}{3}a\right)+(1+a)z-2a\dfrac{z^3}{3}, 当-1<z\leqslant 0, \\ \left(1+\dfrac{1}{3}a\right)-(1+a)z+2a\dfrac{z^3}{3}，当0<z<1, \\ 0\qquad ，其他。\end{cases}$$

显然ζ是对称的。

(*)3—11 ξ、η服从正态分布，但ξ、η不独立，则$\xi+\eta$不一定服从正态分布的例子

有关概念

若随机变量ξ、η独立且都服从正态分布，则$\xi+\eta$亦服从正态分布。

但当独立性条件去掉时，上述结论不一定成立。

反例

设f和g是密度函数，F和G是其对应的分布函数．设

$$h(x,\ y)=f(x)g(y)[1+\alpha(2F(x)-1)(2F(y)-1)],$$

其中$|\alpha|\leqslant 1$．

则可验证$h(x,\ y)$是二维密度函数，且$f(x)$、$g(y)$分别是其边际密度．

特别取f和g为$N(0,\ 1)$，即

$$f(x)=g(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2},\quad -\infty<x<\infty. \tag{1}$$

且设 $(\xi,\ \eta)\sim h(x,\ y)$

$$=f(x)g(y)[1+\alpha(2F(x)-1)(2F(y)-1)]$$

则$\xi+\eta$不再是正态分布了（除非$\alpha=0$，此时ξ、η就独立了）．

事实上，若设$\zeta=\xi+\eta$，则

$$E\zeta=0,\ D(\zeta)=D\xi+D\eta+2cov(\xi,\ \eta),$$

易证 $cov(\xi,\ \eta)=\frac{\alpha}{\pi}$．

所以 $D\zeta=2\left[1+\frac{\alpha}{\pi}\right]$．

下面我们来证明$\zeta=\xi+\eta$不服从正态分布．

采用反证法：

设ζ服从正态分布，则ζ的矩母系数为

$$M_\zeta(t)=e^{t^2\left(1+\frac{\alpha}{\pi}\right)}.$$

又由(1)，我们有ζ的矩母系数为

$$M_1(t)=[e^{t^2/2}]^2+\alpha\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{tx+ty}[2F(x)-1]$$

$$\cdot[2F(y)-1]f(x)g(y)dxdy$$

$$=e^{t^2}+\alpha\left[\int_{-\infty}^{\infty}e^{tx}[2F(x)-1]f(x)dx\right]^2,$$

其中 $\int_{-\infty}^{\infty} e^{tx}[2F(x)-1]f(x)dx$

$$= -2\int_{-\infty}^{\infty} e^{tx}[1-F(x)]f(x)dx + e^{\frac{t^2}{2}}$$

$$= e^{\frac{t^2}{2}} - 2\int_{-\infty}^{\infty}\int_{x}^{\infty}\frac{1}{2\pi} exp\left\{-\frac{1}{2}(x^2+u^2-2tx)\right\}dudx$$

$$= e^{\frac{t^2}{2}} - 2e^{\frac{t^2}{2}}p\left\{z_1 > \frac{t}{\sqrt{2}}\right\},$$

其中z_1是$N(0,\ 1)$变量。

所以 $$M_1(t) = e^{t^2} + \alpha[e^{\frac{t^2}{2}} - 2e^{\frac{t^2}{2}}p\left\{z_1 > \frac{t}{\sqrt{2}}\right\}]^2$$

$$= e^{t^2}\left[1+\alpha\left(2-p\left\{z_1 > \frac{t}{\sqrt{2}}\right\}\right)^2\right].$$

若要使 $M_\zeta(t) = M_1(t)$，对$\forall t$和$|\alpha| \leqslant 1$成立。即有

$$e^{t^2}e^{\frac{\alpha}{\pi}t^2} = e^{t^2}\left[1+\alpha\left(1-2p\left\{z_1 > \frac{t}{\sqrt{2}}\right\}\right)^2\right],$$

$\forall t$和$|\alpha| \leqslant 1$。

但当 $\alpha > 0$ 时，上式右边括号内的表达式是有界的，其上界为 $1+\alpha$，然而$e^{\frac{\alpha}{\pi}t^2}$ 是无界的。所以上面等式对$\forall t$和$|\alpha| \leqslant 1$ 不可能成立。

即$\zeta = \xi + \eta$不服从正态分布。

3—12 两个不同的联合分布(函数)，它们可以有相同的边际分布的例子

有关概念

设二维随机向量$(\xi_1,\ \xi_2)$的联合分布函数为 $F(x,\ y)$，称

$F_\xi(x)=F(x,\ \infty)$ 为随机变量 ξ 的边际分布函数，$F_\eta(y)=F(\infty,\ y)$为随机变量η的边际分布函数。

(1) 对二维离散型随机向量$(\xi_1,\ \xi_2)$，设ξ_1取值$x_1,\ x_2,\ \cdots$；ξ_2取值$y_1,\ y_2,\ \cdots$；则

$$p(x_i,\ y_j)=p(\xi_1=x_i,\ \xi_2=y_j)$$
$$(i,\ j=1,\ 2,\ \cdots)$$

称为$(\xi_1,\ \xi_2)$的分布列。此时ξ_1与ξ_2的边际分布列为：

$$p_i=\sum_j p(x_i,\ y_j)=p(\xi_1=x_i)\quad (i=1,\ 2,\ \cdots),$$

$$p_j=\sum_i p(x_i,\ y_j)=p(\xi_2=y_j)\quad (j=1,\ 2,\ \cdots).$$

(2) 对二维连续型随机向量，设其联合密度函数为 $f(x,\ y)$，则ξ_1和ξ_2的边际密度函数分别为

$$f_{\xi_1}(x)=\int_{-\infty}^{\infty}f(x,\ y)dy$$

和
$$f_{\xi_2}(y)=\int_{-\infty}^{\infty}f(x,\ y)dx.$$

我们知道，由联合分布能够确定边际分布，但联合分布不能由边际分布唯一确定。最典型的例子当然应该是二元正态分布。对于任何满足条件$|r|<1$的相关系数r且各变元均值，方差相同，则它们的边际分布总是一样的（由读者自行验证）。下面我们再来看两个随机变量ξ_1和ξ_2的（边际）分布相同，但它们的联合分布却完全不同的例子。

反例

例1 设袋中装有2只白球和3只黑球，现用两种方法摸球：(1) 有放回地摸球，(2) 无放回地摸球。

定义下列随机变量：

$$\xi_1=\begin{cases}1, & \text{第一次摸出白球,}\\0, & \text{第一次摸出黑球.}\end{cases}$$

$$\xi_2=\begin{cases}1, & \text{第二次摸出白球,}\\0, & \text{第二次摸出黑球.}\end{cases}$$

在第(1)种情况下，采用有放回的摸球，则(ξ_1, ξ_2)的联合概率分布和边际分布由(表一)给出.

ξ_2/ξ_1	0	1	$p(\xi_2=y_j)$
0	$\frac{3}{5}\cdot\frac{3}{5}$	$\frac{2}{5}\cdot\frac{3}{5}$	$\frac{3}{5}$
1	$\frac{3}{5}\cdot\frac{2}{5}$	$\frac{2}{5}\cdot\frac{2}{5}$	$\frac{2}{5}$
$p(\xi_1=x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	

(表一)

第(2)种情况采用不放回摸球，则(ξ_1, ξ_2)的联合概率分布和边际分布由(表二)给出.

ξ_2/ξ_1	0	1	$p(\xi_2=y_j)$
0	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{1}{4}$	$\frac{2}{5}$
$p(\xi_1=x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	

(表二)

这个例子表明，两种情况下ξ_1和ξ_2的（边际）分布是相同

的，但它们的联合分布却不同。

例2 设两个二元分布函数为$F(x, y)$及$G(x, y)$，分别有密度函数为

$$f(x, y)=\begin{cases} x+y, & 若 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1; \\ 0, & 其他. \end{cases}$$

$$g(x, y)=\begin{cases} \left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right), & 若 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1, \\ 0, & 其他. \end{cases}$$

容易看出，方程

$$\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)=x+y$$

只有根为$x=\frac{1}{2}$，$y=\frac{1}{2}$，因而在正方形$0\leqslant x\leqslant 1$，$0\leqslant y\leqslant 1$中，只是沿两直线$x=\frac{1}{2}$及$y=\frac{1}{2}$上，才有$f(x, y)=g(x, y)$，可见$F(x, y)$与$G(x, y)$不恒等。然而，两个（边际）分布函数恒等，这是因为它们的两个密度函数相等的缘故。

事实上，

$$\int_{-\infty}^{\infty} f(x, y)dy=\int_0^1 (x+y)dy$$

$$=\frac{1}{2}+x=\int_0^1\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)dy$$

$$=\int_{-\infty}^{\infty} g(x, y)dy\quad(0\leqslant x\leqslant 1),$$

$$\int_{-\infty}^{\infty} f(x, y)dx=\int_0^1 (x+y)dx=\frac{1}{2}+y$$

$$=\int_0^1\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)dx=\int_{-\infty}^{\infty} g(x, y)dx\qquad(0\leqslant y\leqslant 1).$$

注　文献[6]P160EX19也说明对于一个给定的边际分布密度，可以有无穷多个联合分布密度f_α与之对应。这个著名例子是法国的E. J. Gumbel于1958年构造的。

3—13 随机变量ξ_1，ξ_2，ξ_3，两两独立，但不相互独立的例子

有关概念和命题

1 定义 设ξ_1，ξ_2，…，ξ_n为n个随机变量，若对任意的实数x_1，x_2，…，x_n，都有

$$p\{\xi_1<x_1,\ \xi_2<x_2,\ \cdots,\ \xi_n<x_n\}$$
$$=p(\xi_1<x_1)p(\xi_2<x_2)\cdots p(\xi_n<x_n), \qquad (*)$$

则称ξ_1，ξ_2，…，ξ_n是相互独立的。

2 若ξ_i的分布函数为$F_i(x)$，它们的联合分布函数为$F(x_1, x_2, \cdots, x_n)$，则(＊)等价于对一切x_1，…，x_n有

$$F(x_1,\ x_2,\ \cdots,\ x_n)=F_1(x_1)\cdots F_n(x_n). \qquad (*-1)$$

3 对离散型随机变量，(＊)等价于对任何一组可能取的值$(x_1, x_2, \cdots, x_n)$有

$$p(\xi_1=x_1,\ \xi_2=x_2,\ \cdots,\ \xi_n=x_n)$$
$$=p(\xi_1=x_1)p(\xi_2=x_2)\cdots p(\xi_n=x_n). \qquad (*-2)$$

4 对连续型随机变量，(＊)的等价形式是对一切$x_1, x_2, \cdots, x_n$有

$$f(x_1,\ x_2,\ \cdots,\ x_n)=f_1(x_1)f(x_2)\cdots f(x_n) \qquad (*-3)$$

几乎处处成立（关于R_n中的勒贝格测度）。其中$f_i(x_i)$是ξ_i的密度函数，而$f(x_1, x_2, \cdots, x_n)$是联合分布密度函数。若f和诸f_i都是连续函数，则（＊）成立的充要条件是(＊—3）处处成立。这是因为二个连续函数如果几乎处处相等，则必恒等。

首先我们注意到有如下的

定理 若随机变量整体独立，则随机变量的一部分也相互

独立．

即若ξ_1，ξ_2，…，ξ_n相互独立，则对任意K（K是正整数且$2\leqslant k\leqslant n$）个随机变量ξ_{i1}，ξ_{i2}，…，ξ_{ik}它们也相互独立。

证明 不失一般性，可设$i_1=1$，$i_2=2$，…，$i_k=k$。并记$(\xi_1,\cdots,\xi_k)$的联合分布函数为

$$F_{\xi_1,\cdots,\xi_k}(x_1, x_2, \cdots, x_k),$$

则由ξ_1，ξ_2，…，ξ_n的相互独立性，可得

$$F_{\xi_1,\cdots,\xi_k}(x_1, x_2, \cdots, x_k)$$

$$= \lim_{\substack{x_{k+1}\to\infty\\ \vdots \\ x_n\to\infty}} F(x_1, x_2, \cdots, x_k, x_{k+1}, \cdots, x_n)$$

$$= \lim_{\substack{x_{k+1}\to\infty\\ \vdots \\ x_n\to\infty}} F_{\xi_1}(x_1)F_{\xi_2}(x_2)\cdots F_{\xi_k}(x_k)F_{\xi_{k+1}}(x_{k+1})\cdots F_{\xi_n}(x_n)$$

$$= F_{\xi_1}(x_1)\cdots F_{\xi_k}(x_k),$$ 证毕。

但定理的逆不真，我们有如下的

反例

例1（连续型随机变量的例子） 设(ξ_1, ξ_2, ξ_3)的联合分布密度为

$$f(x, y, z)=\begin{cases}\dfrac{1}{8\pi^3}(1-\sin x\sin y\sin z), & \text{当}0\leqslant x\leqslant 2\pi,\\ & 0\leqslant y\leqslant 2\pi,\\ & 0\leqslant z\leqslant 2\pi,\\ 0, & \text{其他。}\end{cases}$$

则当 $0\leqslant x\leqslant 2\pi$，$0\leqslant y\leqslant 2\pi$ 时，ξ_1和ξ_2的联合分布密度函数为

$$f_{\xi_1,\xi_2}(x, y)=\int_0^{2\pi}\frac{1}{8\pi^3}(1-\sin x\sin y\sin z)dz$$

$$=\left[\frac{z}{8\pi^3}+\sin x\sin y\cos z\right]_0^{2\pi}$$

$$=\frac{1}{4\pi^2},$$

故有 $f_{\xi_1,\xi_2}(x, y)=\begin{cases}\frac{1}{4\pi^2}, & 0\leqslant x\leqslant 2\pi,\ 0\leqslant y\leqslant 2\pi,\\ 0, & \text{其他。}\end{cases}$

当$0\leqslant x\leqslant 2\pi$时，$\xi_1$的密度函数

$$f_{\xi_1}(x)=\int_{-\infty}^{\infty} f_{\xi_1,\xi_2}(x, y)dy$$

$$=\int_0^{2\pi}\frac{1}{4\pi^2}dy=\frac{1}{2\pi}.$$

当$x\overline{\in}[0, 2\pi]$时，$f_{\xi_1}(x)=0$。

所以

$$f_{\xi_1}(x)=\begin{cases}\frac{1}{2\pi}, & x\in[0, 2\pi],\\ 0, & x\overline{\in}[0, 2\pi].\end{cases}$$

类似地可得ξ_2的密度函数为

$$f_{\xi_2}(y)=\begin{cases}\frac{1}{2\pi}, & y\in[0, 2\pi],\\ 0, & y\overline{\in}[0, 2\pi].\end{cases}$$

由此可见，对任何$x\in R_1$，$y\in R_1$，都有

$$f_{\xi_1,\xi_2}(x, y)=f_{\xi_1}(x)\cdot f_{\xi_2}(y).$$

注意到ξ_1，ξ_2，ξ_3三者在其联合密度函数中所处的地位相同，所以应该有

$$f_{\xi_1,\xi_3}(x, z)=f_{\xi_1}(x)\cdot f_{\xi_3}(z)\quad(\forall x\in R_1, z\in R_1),$$

$$f_{\xi_2,\xi_3}(y, z)=f_{\xi_2}(y)\cdot f_{\xi_3}(z)\quad(\forall y\in R_1, z\in R_1).$$

所以ξ_1，ξ_2，ξ_3两两独立。

但当$x\in[0, 2\pi]$，$y\in[0, 2\pi]$，$z\in[0, 2\pi]$时

$$f_{\xi_1,\xi_2,\xi_3}(x, y, z) = \frac{1}{8\pi^3}(1 - \sin x \sin y \sin z)$$

$$\neq \frac{1}{8\pi^3} = f_{\xi_1}(x) \cdot f_{\xi_2}(y) \cdot f_{\xi_3}(z).$$

故ξ_1，ξ_2，ξ_3不相互独立。

例2（离散型随机变量的例子） 设(ξ_1, ξ_2, ξ_3)的联合概率分布由下式给出。

$$p(\xi_1 = x, \xi_2 = y, \xi_3 = z)$$

$$= \begin{cases} \dfrac{3}{16}, & 若(x, y, z) \in \left\{ \begin{matrix} (0, 0, 0), & (0, 1, 1) \\ (1, 0, 1), & (1, 1, 0) \end{matrix} \right\}, \\ \dfrac{1}{16}, & 若(x, y, z) \in \left\{ \begin{matrix} (0, 0, 1), & (0, 1, 0) \\ (1, 0, 0), & (1, 1, 1) \end{matrix} \right\}. \end{cases}$$

由此可得

$$p\{\xi_1 = x, \xi_2 = y\}$$

$$= \sum_{z=0,1} p(\xi_1 = x, \xi_2 = y, \xi_3 = z)$$

$$= \frac{1}{4}, \text{ 当}(x, y) \in \{(0,0), (0,1), (1,0), (1,1)\}.$$

类似地求得

$$p(\xi_1 = x, \xi_3 = z) = \frac{1}{4},$$

当 $(x, y) \in \{(0,0), (0,1), (1,0), (1,1)\}$.

$$p(\xi_1 = x, \xi_3 = z) = \frac{1}{4},$$

当$(y,z) \in \{(0,0), (0,1), (1,0), (1,1)\}$.

又有 $p(\xi_1 = x) = \sum\limits_{y=0;1} p(\xi_1 = x, \xi_2 = y) = \frac{1}{2}$，当$x \in \{0,1\}$，

$$p(\xi_2 = y) = \frac{1}{2} \quad 当 y \in \{0, 1\}.$$

$$p(\xi_3 = z) = \frac{1}{2}, \ 当 z \in \{0, 1\}.$$

当 $(x, y, z) = (0, 0, 0)$ 时

$$p(\xi_1 = 0, \ \xi_2 = 0, \ \xi_3 = 0) = \frac{3}{16}$$

$$\neq p(\xi_1 = 0) \cdot p(\xi_2 = 0) \cdot p(\xi_3 = 0)$$

$$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}.$$

但对任何 $x \in \{0, 1\}$，$y \in \{0, 1\}$，$z \in \{0, 1\}$，都有

$$p(\xi_1 = x, \ \xi_2 = y) = p(\xi_1 = x) \cdot p(\xi_2 = y),$$

$$p(\xi_1 = x, \ \xi_3 = z) = p(\xi_1 = x) \cdot p(\xi_3 = z),$$

$$p(\xi_2 = y, \ \xi_3 = z) = p(\xi_2 = y) \cdot p(\xi_3 = z).$$

上述表明 ξ_1，ξ_2，ξ_3 两两独立，但 ξ_1，ξ_2，ξ_3 不相互独立。

3—14 ξ、η 不独立，但 ξ^2 和 η^2 独立的例子

有关命题

定理 设 ξ 和 η 是独立的随机变量，并且 f 和 g 是Borel函数，则 $f(\xi)$ 和 $g(\eta)$ 也是相互独立的。

但在 $f(\xi)$ 和 $g(\eta)$ 相互独立的条件下，未必能推出 ξ 和 η 之间的独立性。

反例

例1 设 ξ 和 η 的联合分布密度函数为

注：请参看例2—3的“进一步的讨论”。

$$f(x, y)=\begin{cases}\dfrac{1+xy}{4}, & |x|<1,\ |y|<1,\\ 0, & \text{其他}.\end{cases}$$

因为ξ和η的边际分布密度分别为

$$f_\xi(x)=\int_{-1}^{1}\frac{1+xy}{4}dy=\frac{1}{2}\quad(|x|<1)$$

和 $$f_\eta(y)=\frac{1}{2}\quad(|y|<1).$$

当$|x|<1$，$|y|<1$时

$$f(x, y)\neq f_\xi(x)\cdot f_\eta(y),$$

所以ξ和η不独立.

然而ξ^2和η^2却是相互独立的. 下面就来证明这一点. 先求ξ^2的分布函数，当$0<x\leqslant1$时

$$F_{\xi^2}(x)=p(\xi^2<x)=\int_{t^2<x}f_\xi(t)dt$$
$$=\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{2}dt=\sqrt{x}.$$

$x\leqslant0$时，$F_{\xi^2}(x)=0$，$x>1$时，$F_{\xi^2}(x)=1$，同理可证

$$F_{\eta^2}(y)=\begin{cases}0, & y\leqslant0,\\ \sqrt{y}, & 0<y\leqslant1,\\ 1, & y>1.\end{cases}$$

把(ξ^2, η^2)的联合分布函数记为$F^*(x, y)$，则

$$F^*(x, y)=p(\xi^2<x, \eta^2<y).$$

$$=\begin{cases}1 & x>1,\ y>1,\\ p(\eta^2<y)=\sqrt{y} & x>1,\ 0<y\leqslant1,\\ p(\xi^2<x)=\sqrt{x} & 0<x\leqslant1,\ y>1,\\ \displaystyle\iint\limits_{\left(\substack{t^2<x\\ s^2<y}\right)}\frac{1+st}{4}dsdt=\sqrt{xy} & 0<x\leqslant1,\ 0<y\leqslant1,\\ 0 & x\leqslant0\text{或}y\leqslant0.\end{cases}$$

不难验证

$$F^{*}(x, y)=F_{\xi^2}(x)\cdot F_{\eta^2}(y).$$

对所有x，y都成立，所以ξ^2和η^2相互独立。

例2 设

$$p(\xi=-1, \eta=-1)=p(\xi=-1, \eta=1)$$

$$=p(\xi=1, \eta=-1)=\frac{6}{20},$$

$$p(\xi=1, \eta=1)=\frac{2}{20}.$$

则显然ξ与η不独立，但

$$\xi^2\equiv 1, \eta^2\equiv 1.$$

故ξ^2与η^2相互独立。

进一步的讨论

注意 由ξ和η独立，可以推出ξ与η^2独立，但是由ξ与η^2独立并不能推出ξ与η独立（在某些读者和著作中却有时使用ξ与η^2独立则ξ与η也独立的错误结论）。

现举反例如下：

设（ξ，η）的联合分布密度为

$$f(x, y)=\begin{cases}\dfrac{1+xy}{4}, & \text{当}|x|<1,\ |y|<1,\\ 0, & \text{其他}.\end{cases}$$

则ξ与η不独立，已于上例阐述，然而此处ξ与η^2却是独立的。事实上，ξ之分布函数为

$$F_{\xi}(x)=p(\xi<x)=\int_{-1}^{x}\frac{1}{2}dt$$

$$
=\begin{cases}\frac{1}{2}(1+x), & 当|x|<1时,\\ 0, & 当x\leqslant -1时,\\ 1, & x\geqslant 1.\end{cases}
$$

η^2之分布函数为

$$
F_{\eta^2}(y)=\begin{cases}0, & 当y\leqslant 0,\\ \sqrt{y}, & 当0<y\leqslant 1,\\ 1, & 当y>1.\end{cases}
$$

$(\xi,\ \eta^2)$ 的联合分布函数为

$$
F_{\xi,\eta^2}(x,\ y)=p(\xi<x,\ \eta^2<y)
$$

$$
=\begin{cases}p(\xi<x)=\frac{1}{2}(1+x), & 当|x|<1,y>1时,\\ p(\eta^2<y)=\sqrt{y}, & 当x\geqslant 1,0<y\leqslant 1时,\\ 0, & 当x\leqslant -1或y\leqslant 0时,\\ 1, & 当x\geqslant 1,\ y>1时,\\ \int_{-1}^{x}ds\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1+st}{4}dt=\left(\frac{1+x}{2}\right)\sqrt{y}, & 当|x|<1,0<y\leqslant 1时.\end{cases}
$$

由此知 $F_{\xi,\eta^2}(x,\ y)=F_{\xi}(x)F_{\eta^2}(y)$.

对所有的x、y都成立。由此我们得到ξ与η^2独立，但ξ与η不独立的反例。

3—15 相同的随机向量构造的不同的Borel可测函数（不恒等于常数）之间也可能是独立的例子

反例

设ξ，η为独立随机变量且具有相同的指数分布密度函数

$$f(x)=\begin{cases}e^{-x}, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

对随机向量（ξ，η）构造下列函数：

$$\alpha=\alpha(\xi,\ \eta)=\xi+\eta,$$

$$\beta=\beta(\xi,\ \eta)=\xi/\eta.$$

下证α与β相互独立。

先求α与β的联合分布函数$F_{\alpha,\beta}(u,\ v)$。由于ξ和η的联合分布密度函数为

$$f_{\xi,\eta}(x,\ y)=\begin{cases}e^{-(x+y)}, & x\geqslant 0,\ y\geqslant 0,\\ 0, & 其他\end{cases}$$

因此 $$F_{\alpha,\beta}(u,\ v)=p(\alpha<u,\ \beta<v)$$

$$=p(\xi+\eta<u,\ \xi/\eta<v)$$

$$=\iint\limits_{\begin{Bmatrix}x+y<u\\ x/y<v\\ x\geqslant 0\\ y\geqslant 0\end{Bmatrix}} f_{\xi,\eta}(x,\ y)dxdy=\iint\limits_{\begin{Bmatrix}x+y<u\\ x/y<v\\ x\geqslant 0\\ y\geqslant 0\end{Bmatrix}} e^{-(x+y)}dxdy.$$

对$u\geqslant 0$，$v\geqslant 0$作变换

$$\begin{cases}u=x+y,\\ v=\dfrac{x}{y}.\end{cases}$$

反函数为

$$\begin{cases}y=\dfrac{u}{v+1},\\ x=\dfrac{uv}{1+v}.\end{cases}$$

雅可比行列式

$$J=\frac{\partial(x,\ y)}{\partial(u,\ v)}=\begin{vmatrix}\frac{v}{1+v}, & \frac{u}{(1+v)^2}\\ \frac{1}{1+v}, & -\frac{u}{(1+v)^2}\end{vmatrix}=\frac{-u}{(1+v)^2},$$

所以α与β的联合密度函数为

$$f_{\alpha,\beta}(u,\ v)=f_{\xi,\eta}(x(u,\ v),\ y(u,\ v))|J|$$

$$=e^{-(x+y)}\ |J|=e^{-\left(\frac{uv}{1+v}+\frac{u}{1+v}\right)}\cdot\frac{u}{(1+v)^2}$$

$$=e^{-u}\cdot\frac{u}{(1+v)^2}\quad(u\geqslant 0,\ v\geqslant 0).$$

再求 $$f_\alpha(u)=\int_0^{+\infty}f_{\alpha,\beta}(u,\ v)dv$$

$$=ue^{-u}\left(-\frac{1}{(1+v)}\right)_0^{\infty}=ue^{-u}\ (u\geqslant 0).$$

$$f_\beta(v)=\int_0^{+\infty}f_{\alpha,\beta}(u,\ v)du$$

$$=\frac{1}{(1+v)^2}\int_0^{\infty}ue^{-u}du$$

$$=\frac{1}{(1+v)^2}\left[-ue^{-u}\Big|_0^{\infty}+\int_0^{\infty}e^{-u}du\right]$$

$$=\frac{1}{(1+v)^2}\qquad(v\geqslant 0),$$

得 $f_{\alpha,\beta}(u,\ v)=f_\alpha(u)\cdot f_\beta(v)$ （对一切$(u,\ v)$）都成立，所以α与β相互独立。

3—16 从随机向量$\xi=(\xi_1,\xi_2,\cdots,\xi_m)$和$\eta=(\eta_1,\eta_2,\cdots,\eta_n)$之间的独立性推不出$\xi$的分量$\xi_1,\xi_2,\cdots,\xi_m$或者$\eta$的分量$\eta_1,\eta_2,\cdots,\eta_n$之间的独立性的例子

有关概念和命题

定义 两个随机向量$\xi=(\xi_1, \xi_2, \cdots, \xi_m)$与$\eta=(\eta_1, \eta_2, \cdots, \eta_n)$称为独立的，如果

$$F(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n)$$
$$=F_1(x_1, x_2, \cdots, x_m)\cdot F_2(y_1, y_2, \cdots, y_n)$$

对所有$(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n)\in R_{m+n}$都成立。这里的$F$、$F_1$、$F_2$分别是$(\xi_1, \xi_2, \cdots, \xi_m, \eta_1, \cdots, \eta_n)$，$(\xi_1, \xi_2, \cdots, \xi_m)$和$(\eta_1, \eta_2, \cdots, \eta_n)$的联合分布函数。

显然，若随机向量ξ和η独立，则ξ的子向量也是独立的。但是，从随机向量ξ和η之间的独立性推不出ξ的分量或者η的分量之间的独立性。

反例

设随机向量$\xi=(\xi_1, \xi_2)$具有联合密度函数

$$f_{\xi_1,\xi_2}(x,y)=\begin{cases}\frac{1}{4}[1+xy(x^2-y^2)], & |x|\leqslant 1,\ |y|\leqslant 1,\\ 0, & \text{其他.}\end{cases}$$

随机向量$\eta=(\eta_1, \eta_2)$具有联合密度函数

$$g_{\eta_1,\eta_2}(u,v)=\frac{1}{2\pi\sqrt{1-p^2}}exp\cdot\left\{-\frac{1}{2(1-p^2)}(u^2-2puv+v^2)\right\}$$
$$(-\infty<u<\infty,\ -\infty<v<+\infty).$$

其中$|p|<1$是一给定常数。

设$(\xi_1, \xi_2, \eta_1, \eta_2)$的联合密度函数由下式给出：

$$h(x, y, u, v)$$

$$= \frac{1}{8\pi\sqrt{1-p^2}}[1+xy(x^2-y^2)]$$

$$\cdot exp\left\{-\frac{1}{2(1-p^2)}(u^2-2puv+v^2)\right\}$$

$$(|x|\leqslant 1, \quad |y|\leqslant 1, \quad -\infty<u, \quad v<\infty)$$

则显然(ξ_1, ξ_2)和(η_1, η_2)是独立的，但ξ_1与ξ_2或者η_1与η_2不是独立的。事实上，ξ_1的密度函数

$$f_{\xi_1}(x)=\int_{-1}^{1}\frac{1}{4}[1+xy(x^2-y^2)]dy$$

$$=\frac{1}{4}\left[y+\frac{x^3y^2}{2}-\frac{xy^4}{4}\right]_{-1}^{1}$$

$$=\frac{1}{2} \qquad (|x|\leqslant 1),$$

$$f_{\xi_1}(x)=0 \qquad (|x|>1).$$

ξ_2的密度函数

$$f_{\xi_2}(y)=\begin{cases}\frac{1}{4}\int_{-1}^{1}[1+xy(x^2-y^2)]dx=\frac{1}{2}, & (|y|\leqslant 1),\\ 0, & (|y|>1).\end{cases}$$

显然，当$|x|\leqslant 1$，$|y|\leqslant 1$，且$|x|\neq|y|$时，

$$f_{\xi_1,\xi_2}(x, y)\neq f_{\xi_1}(x)\cdot f_{\xi_2}(y),$$

所以ξ_1和ξ_2不相互独立。

类似地可证η_1和η_2不相互独立。因为η_1和η_2的密度函数分别为

$$g_{\eta_1}(u)=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \qquad (-\infty<u<+\infty),$$

$$g_{\eta_2}(v)=\frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}} \qquad (-\infty<v<+\infty).$$

当$\rho\neq0$时

$$g_{\eta_1,\eta_2}(u,\ v)\neq g_{\eta_1}(u)\cdot g_{\eta_2}(v).$$

3—17 ξ和η_1独立，ξ和η_2独立，但ξ和随机向量$(\eta_1,\ \eta_2)$不独立的例子

有关概念

见例3—15。

反例

设随机变量η_1，η_2相互独立，且皆以概率$\frac{1}{2}$取值$+1$和-1。令$\xi=\eta_1\cdot\eta_2$，则ξ和η_1之间，ξ和η_2之间均独立，但ξ与随机向量（η_1，η_2）不独立。

先证ξ和η_1，ξ与η_2均独立．由假设知

$$\eta_1:\begin{pmatrix}1 & -1\\ 1/2 & 1/2\end{pmatrix} \qquad \eta_2:\begin{pmatrix}1 & -1\\ 1/2 & 1/2\end{pmatrix}$$

且η_1和η_2相互独立．由此可得

$$\begin{aligned}P(\xi=1)&=P(\{\eta_1=1,\eta_2=1\}\cup\{\eta_1=-1,\eta_2=-1\})\\&=P(\eta_1=1,\eta_2=1)+P(\eta_1=-1,\eta_2=-1)\\&=P(\eta_1=1)\cdot P(\eta_2=1)+P(\eta_1=-1\cdot P(\eta_2=-1)\\&=1/2\cdot1/2+1/2\cdot1/2=1/2,\end{aligned}$$

$$\begin{aligned}P(\xi=-1)&=P(\{\eta_1=1,\eta_2=-1\}\cup\{\eta_1=-1,\eta_2=1\})\\&=1/2\cdot1/2+1/2\cdot1/2=1/2,\end{aligned}$$

$P(\xi=1,\ \eta_1=1)$

$$=P([(\eta_1=1,\eta_2=1)\cup(\eta_1=-1,\eta_2=-1)]$$
$$\cap(\eta_1=1))$$
$$=P(\eta_1=1,\eta_2=1)=P(\eta_1=1)\cdot P(\eta_2=1)$$
$$=1/4=P(\eta_1=1)P(\xi=1).$$

$$P(\xi=-1,\ \eta_1=1)$$
$$=P([(\eta_1=1,\eta_2=-1)\cup(\eta_1=-1,\eta_2=1)]$$
$$\cap(\eta_1=1))$$
$$=P(\eta_1=1,\eta_2=-1)$$
$$=P(\eta_1=1)\cdot P(\eta_2=-1)$$
$$=1/2\cdot 1/2=1/4$$
$$=P(\xi=-1)P(\eta_1=1).$$

类似地可证

$$P(\xi=1,\eta_1=-1)=P(\xi=1)\cdot P(\eta_1=-1),$$
$$P(\xi=-1,\eta_1=-1)=P(\xi=-1)\cdot P(\eta_1=-1).$$

上述表明ξ与η_1独立．用同样的方法可证，ξ与η_2也独立．但是，ξ与$(\eta_1,\ \eta_2)$不独立．事实上，因为

$$\{(\eta_1,\eta_2)=(1,1)\}\subset\{\xi=1\},$$

所以 $$P(\xi=1,\ (\eta_1,\eta_2)=(1,1))$$
$$=P((\eta_1,\eta_2)=(1,1))=\frac{1}{4}.$$

但是 $$P(\xi=1)\cdot P((\eta_1,\eta_2)=(1,1))$$
$$=P(\xi=1)P(\eta_1=1)P(\eta_2=1)=\frac{1}{8},$$

所以 $$P((\xi=1),\ (\eta_1,\ \eta_2)=(1,\ 1))$$
$$\neq P(\xi=1)\cdot P((\eta_1,\ \eta_2)=(1,\ 1)).$$

由此推出ξ与$(\eta_1,\ \eta_2)$不独立．

(*) 3—18 非独立随机变量序列的例子

有关概念

设ξ_1，ξ_2，…，ξ_n…为随机变量序列，若对任意$k(k=2, 3, \cdots, n, \cdots)$个随机变量$\xi_{i_1}$，$\xi_{i_2}$，…，$\xi_{i_k}$，$\{i_1, i_2, \cdots, i_k\}\subset\{1, 2, 3, \cdots, n, \cdots\}$，它们的联合分布函数

$$F_{i_1, i_2, \cdots, i_k}(x_{i_1}, x_{i_2}, \cdots, x_{i_k}) = F_{i_1}(x_{i_1})F_{i_2}(x_{i_2})\cdots F_{i_k}(x_{i_k}). \quad (*)$$

其中$F_{i_r}(r=1, 2, \cdots, k)$为$\xi_{i_r}$的分布函数，$x_{i_r}\in R_1(r=1, 2, \cdots, k)$。则称随机变量序列$\{\xi_n\}$为相互独立的随机变量序列。

如果ξ_1，ξ_2，…，ξ_n，…都是离散型随机变量，且ξ_i的分布集中在集A_i上$(i=1, 2, \cdots, n, \cdots)$，则等式(*)等价于

$$P(\xi_{i_1}=a_{i_1}, \xi_{i_2}=a_{i_2}, \cdots, \xi_{i_k}=a_{i_k}) = P(\xi_{i_1}=a_{i_1})\cdot P(\xi_{i_2}=a_{i_2})\cdots P(\xi_{i_k}=a_{i_k}). \quad (*-1)$$

对任意$a_{i_m}\in A_{i_m}(m=1, 2, \cdots, k)$成立。

反例

在掷硬币的独立试验序列里，考虑“国徽与字出现的累积次数相等”这一事件E。以E为根据引进一个随机变量N_r，它表示r个试验里E的出现次数。可以证明N_r的分布律收敛于正态律（见‹13›§11.4）。但N_r并不像 Lindeberg–Ляпунов 定理中那样是独立随机变量的和：$\xi_1+\xi_2+\cdots+\xi_r$。这里的ξ_k随E在第k次试验出现或不出现而取值1或0。由N_r的意义可知，如果$\xi_k=1$，则ξ_{k+1}决不能等于1；$\xi_k=0$的时候，ξ_{k+1}可以取值0，也许取值1。因为

$$P(\xi_{k+1}=1|\xi_k=1)=0,$$

$$P(\xi_{k+1}=0|\xi_k=1)=1,$$

$$\begin{aligned}&P(\xi_k=1,\ \xi_{k+1}=1)\\=&P(\xi_{k+1}=1|\xi_k=1)\cdot P(\xi_k=1)\\=&0\neq P(\xi_k=1)\cdot P(\xi_{k+1}=1).\end{aligned}$$

于是ξ_k和ξ_{k+1}不独立（$k=1, 2, \cdots$），从而$\xi_1, \xi_2, \cdots, \xi_n, \cdots$不是独立随机变量序列.

(*) 3—19 分布函数F_1和F_2的卷积绝对连续，但F_1和F_2不绝对连续的例子

有关定理

首先我们有如下的正面命题：

定理1 设ξ、η独立，它们的分布函数F_1和F_2至少有一个绝对连续，则F_1和F_2的卷积

$$F(x)=\int_{-\infty}^{\infty}F_1(x-y)dF_2(y)$$

亦绝对连续（参阅〔1〕P150），但其逆不真．请看

反例

例1 设$\{\xi_n, n\geqslant 1\}$相互独立相同分布，满足

$$P(\xi_n=\pm 1)=\frac{1}{2}\qquad (n\geqslant 1).$$

作
$$\xi=\sum_{j=1}^{\infty}\frac{\xi_{2j}}{2^{2j}},\qquad \eta=\sum_{j=1}^{\infty}\frac{\xi_{2j-1}}{2^{2j-1}},$$

则ξ与η独立．其相应的特征函数为

$$\varphi_\xi(t)=\prod_{j=1}^{\infty}\cos\frac{t}{2^{2j}},\qquad \varphi_\eta(t)=\prod_{j=1}^{\infty}\cos\frac{t}{2^{2j-1}},$$

而相应的分布函数都产生与L—测度相互奇异的测度，但$\xi+\eta$的特征函数为

$$\varphi_{\xi+\eta}(t)=\frac{\sin t}{t}.$$

相应的分布函数为$[-1, 1]$上的均匀分布，它为绝对连续。

进一步的讨论

当定理1的条件：ξ与η独立不成立时结论不再成立。

例2 设ξ是离散型随机变量，η是连续型随机变量，则

$$\xi=\eta+(\xi-\eta).$$

此时η与$(\xi-\eta)$不独立。虽然η的分布函数绝对连续，但ξ的分布函数不是绝对连续的。

(*) 3—20 随机变量ξ、η的各阶矩不全存在，即使ξ、η已经存在的各阶矩相等，亦不能推出ξ、η的分布相同的例子

众所周知，若ξ、η的分布相同，则ξ、η对应的各阶矩亦相同。但是如果ξ、η的各阶矩不全存在，那么ξ、η即使已经存在的各阶矩相同，也不能推出ξ、η的分布相同。请看

反例

设 $P\left(\xi=\frac{3^k}{k^2}\right)=\frac{2}{3^k}$, $k=1, 2, \cdots$

则 $$E\xi=\sum_{k=1}^{\infty}\frac{2}{k^2}<\infty,$$

且 $$E\xi^2=\sum_{k=1}^{\infty}\frac{2 3^k}{k^4}=\infty.$$

又设 $P(\eta=0)=\frac{1}{3}$, $P\left(\eta=\frac{3^{k+1}}{k^2}\right)=\frac{2}{3^{k+1}}$. $k=1, 2, \cdots$

则 $$E\eta=\sum_{k=1}^{\infty}\frac{2}{k^2}<\infty,$$

且 $$E\xi=E\eta,$$

又 $$E\eta^2=2\sum_{k=1}^{\infty}\frac{3^{k+1}}{k^4}=\infty,$$

但ξ与η的分布不相同。

(*) 3—21 ξ与$\frac{1}{\xi}$有相同的分布，而ξ并非服从哥西分布的例子

有关记号和命题

记 $$\mathscr{C}(\mu, \theta)=\frac{\mu}{\pi\mu^2[1+(x-\theta)^2]},$$

则 $$\mathscr{C}(1, 0)=\frac{1}{\pi(1+x^2)}.$$

首先，我们有如下的

定理 $\xi\sim\mathscr{C}(1, 0)\Longleftrightarrow\frac{1}{\xi}\sim\mathscr{C}(1, 0)$,

证明 "$\Rightarrow$". 设 $\xi\sim\mathscr{C}(1, 0)=\frac{1}{\pi}\frac{1}{1+x^2}$.

令$\eta=\frac{1}{\xi}$，我们来证明

$$\frac{1}{\xi}\sim\frac{1}{\pi}\frac{1}{1+y^2},$$

设$y=\frac{1}{x}$，则$x=\frac{1}{y}$，$x'=-\frac{1}{y^2}$，

$$\therefore \qquad |J|=\frac{1}{y^2}.$$

于是 $$\eta\sim\frac{1}{\pi}\frac{1}{1+\left(\frac{1}{y}\right)^2}\cdot\frac{1}{y^2}=\frac{1}{\pi}\frac{1}{1+y^2},$$

“$\Leftarrow$”，若 $$\eta=\frac{1}{\xi}\sim\frac{1}{\pi}\frac{1}{1+y^2}.$$

设 $y=\frac{1}{x}$，则$y'=-\frac{1}{x^2}$，$|J|=\frac{1}{x^2}$，

$$\therefore \qquad \xi\sim\frac{1}{\pi\left[1+\left(\frac{1}{x}\right)^2\right]}\frac{1}{x^2}=\frac{1}{\pi}\frac{1}{1+x^2},$$

证毕。

然而有趣的是，若ξ与$\frac{1}{\xi}$在$(-\infty,\ +\infty)$上有相同的分布，我们并不能得到$\xi\sim\mathscr{C}(1,\ 0)$的结论，请看

反例

设

$$\xi\sim p(x)=\begin{cases}\frac{1}{4} & \text{当}|x|\leqslant 1,\\ \frac{1}{4x^2} & \text{当}|x|>1.\end{cases}$$

由此不难验证$\frac{1}{\xi}$与ξ有相同的分布。事实上，令$\eta=\frac{1}{\xi}$，且设$y=\frac{1}{x}$，则$x=\frac{1}{y}$，$x'=-\frac{1}{y^2}$，$|J|=\frac{1}{y^2}$，

$$\therefore \qquad \frac{1}{\xi}\sim\begin{cases}\frac{1}{4}\frac{1}{y^2}, & \text{当}\left|\frac{1}{y}\right|\leqslant 1,\ \text{即}|y|\geqslant 1\text{时}\\ \frac{1}{4\left(\frac{1}{y}\right)^2}\frac{1}{y^2}=\frac{1}{4}, & \text{当}\left|\frac{1}{y}\right|>1,\ \text{即}|y|<1\text{时}\end{cases}$$

但是显然ξ并非服从$\mathscr{C}(1,\ 0)$分布。

进一步的讨论

然而当ξ与$\frac{1}{\xi}$具有相同的分布并且这个分布是稳定分布时，则ξ必是$\mathscr{C}(1,\ 0)$分布（可参阅*M. V. Menon*：*Ann Math Stat.* 33(1962) *PP*1267—1271)。

(*)3—22 ξ与1—ξ有相同的分布，但ξ并非服从$B(\alpha,\ \alpha)$分布的例子

有关概念和命题

定义 若ξ之分布密度为

$$P(x)=\begin{cases}\dfrac{1}{B(\alpha,\beta)}x^{\alpha-1}\ (1-x)^{\beta-1}, & \text{当}0<x<1,\\ 0, & \text{其他。}\end{cases}$$

其中 $B(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \qquad \alpha>0,\ \beta>0$

则称ξ服从参数为α、β的Beta分布，记为

$$\xi\sim B(\alpha,\ \beta).$$

我们有下面的

定理 $\xi\sim B(\alpha,\alpha)\Longleftrightarrow 1-\xi\sim B(\alpha,\alpha)$。

证明 "$\Rightarrow$"。令$\eta=1-\xi$，且设$x=1-x$，

$\therefore \quad x=1-x,\ x'=-1,\ |J|=1$。

$$\because \quad \xi\sim\begin{cases}\dfrac{1}{B(\alpha,\alpha)}x^{\alpha-1},\ (1-x)^{\alpha-1} & \text{当}0<x<1,\\ 0, & \text{其他。}\end{cases}$$

$$
\therefore \quad \eta \sim \begin{cases} \dfrac{1}{B(\alpha,\alpha)}(1-y)^{\alpha-1}[1-(1-y)]^{\alpha-1} \\ \quad = \dfrac{1}{B(\alpha,\alpha)} y^{\alpha-1}(1-x)^{\alpha-1}, \\ \qquad\qquad \text{当} 0<y<1 \text{时}, \\ 0, \qquad\qquad \text{其他}. \end{cases}
$$

"$\Leftarrow$"，已知

$$
\eta = 1-\xi \sim \begin{cases} \dfrac{1}{B(\alpha,\alpha)} y^{\alpha-1}(1-y)^{\alpha-1}, & \text{当} 0<y<1, \\ 0, & \text{其他}. \end{cases}
$$

设$y=1-x$，$\therefore y'=-1$，$|J|=1$

$$
\therefore \quad \xi \sim \begin{cases} \dfrac{1}{B(\alpha,\alpha)}(1-x)^{\alpha-1}[1-(1-x)]^{\alpha-1} \\ \quad = \dfrac{1}{B(\alpha,\alpha)} x^{\alpha-1}(1-x)^{\alpha-1}, \\ \qquad\qquad \text{当} 0<x<1, \\ 0, \qquad\qquad \text{其他} \end{cases}
$$

证毕。

然而ξ与$1-\xi$有相同的分布，并不能推出ξ一定服从$B(\alpha,\alpha)$分布，请看

反例 设

$$
\xi \sim P(x) = \frac{1}{B(\alpha,\beta)+B(\beta,\alpha)}[x^{\alpha-1}(1-x)^{\beta-1} + \\ +(1-x)^{\alpha-1}x^{\beta-1}] \qquad (0<x<1).
$$

则易见$P(x)=P(1-x)$，$0<x<1$，而$P(1-x)$恰好是$1-\xi$之分布密度，所以ξ与$1-\xi$有相同的分布，但$P(x)$并非是$B(\alpha,\alpha)$分布。

(*)3—23　ξ与η独立同分布，$\zeta=\frac{\xi}{\eta}\mathscr{C}(1,0)=\frac{1}{\pi(1+x^2)}$但$\xi$、$\mu$并非服从正态分布的例子

有关结论

我们知道，若ξ与η独立同分布，$\xi\sim N(0,1)$，则

$$\zeta=\frac{\xi}{\eta}\sim\mathscr{C}(1,0)=\frac{1}{\pi}\frac{1}{1+x^2}.$$

证明

$$\because\qquad(\xi,\eta)\sim\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}},$$

于是由商的密度公式得

$$\zeta\sim P_\zeta(y)=\int_{-\infty}^{\infty}|x|P(xy,x)dx$$

$$=\int_{-\infty}^{\infty}|x|\frac{1}{2\pi}e^{-\frac{(x^2y^2+x^2)}{2}}dx=\frac{1}{\pi}\int_0^{\infty}xe^{-\frac{x^2}{2}(1+y^2)}dx$$

$$=\frac{1}{\pi}\frac{1}{1+y^2}\int_0^{\infty}d\left[-e^{-\frac{1}{2}x^2(1+y^2)}\right]$$

$$=\frac{1}{\pi}\frac{1}{1+y^2}\left[-e^{-\frac{x^2}{2}(1+y^2)}\right]_{x=0}^{x=\infty}$$

$$=\frac{1}{\pi}\frac{1}{1+y^2},\qquad -\infty<y<+\infty,$$

证毕。

但其逆不成立，请看

反例

设ξ、η独立同分布：

$$\xi \sim P(x)=\frac{\sqrt{2}}{\pi}\frac{1}{1+x^4}, \qquad -\infty<x<+\infty,$$

则不难证明

$$\zeta=\frac{\xi}{\eta}\sim\mathscr{C}(1,0)=\frac{1}{\pi(1+x^2)},$$

事实上，

$$\because \qquad (\xi,\eta)\sim P(x,y)=\frac{2}{\pi^2}\frac{1}{1+x^4}\frac{1}{1+y^4},$$

$$\therefore \qquad \zeta\sim P_\zeta(y)=\int_{-\infty}^{\infty}|x|P(xy,x)dx$$

$$=2\int_0^\infty x\frac{2}{\pi^2}\frac{1}{1+(xy)^4}\frac{1}{1+x^4}dx.$$

令$z=x^2$

$$\frac{4}{\pi^2}\int_0^\infty\frac{1}{1+z^2y^4}\frac{1}{1+z^2}\frac{1}{2}dz$$

$$=\frac{2}{\pi^2}\frac{y^4}{y^4-1}\int_0^\infty\left[\frac{1}{1+z^2y^4}-\frac{1}{(1+z^2)y^4}\right]dz$$

$$=\frac{2}{\pi^2}\frac{y^4}{y^4-1}\left[\frac{1}{y^2}\int_0^\infty\frac{1}{1+(zy^2)^2}d(zy^2)-\frac{1}{y^4}\int_0^\infty\frac{1}{1+z^2}dz\right]$$

$$=\frac{2}{\pi^2}\frac{y^4}{y^4-1}\left[\frac{1}{y^2}t_g^{-1}(zy^2)-\frac{1}{y^4}t_g^{-1}z\right]_0^\infty$$

$$=\frac{2}{\pi^2}\frac{y^4}{y^4-1}\left[\frac{1}{y^2}\frac{\pi}{2}-\frac{1}{y^4}\frac{\pi}{2}\right]$$

$$=\frac{1}{\pi}\frac{1}{1+y^2}.$$

然而ξ、η并非服从正态分布。

(*)3—25 ξ与η独立，分别服从$B(\alpha_1,\beta_1)$及$B(\alpha_2,\beta_2)$分布，又$\xi\eta\sim B(\alpha,\beta)$则$\beta=\beta_1+\beta_2$但只有$\alpha=\alpha_1$或$\alpha=\alpha_2$的例子

有关命题

首先我们有如下的命题：

定理 设ξ、η独立，且$\xi\sim B(\alpha_1,\beta_1)$，$\eta\sim B(\alpha_2\beta_2)$，若$\alpha_1=\alpha_2+\beta_2$，则$\xi\eta\sim B(\alpha_2,\beta_1+\beta_2)$；若$\alpha_2=\alpha_1+\beta_1$，则$\xi\eta\sim B(\alpha_1,\beta_1+\beta_2)$（见〔2〕P215）. 但定理之逆只能部分成立，即我们有如下的结论：若$\xi\eta\sim B(\alpha,\beta)$且$\xi$、$\eta$独立，分别服从$B(\alpha_1,\beta_1)$和$B(\alpha_2,\beta_2)$分布，则有$\beta=\beta_1+\beta_2$，但$\alpha=\alpha_1$或$\alpha=\alpha_2$（详见B. Pamachandaran: Ann Math Stat 33(1962)P1212).

(*)3—25 $\sum_{k=1}^{\infty}\frac{m_k}{k!}s^k(m_k=E\xi^k)$对$s$不绝对收敛$\{m_k\}$仍唯一决定$\xi$之分布函数的例子

有关定理

首先我们有如下的

定理 设$\{m_k\}$是随机变量ξ之矩序列，若$\sum_{k=0}^{\infty}\frac{m_k}{K!}s^k$对某个$s>0$绝对收敛，　　(1)

则$\{m_k\}$唯一决定ξ之分布函数（见〔17〕§5.5). 但条件(1)不是必要的，我们有如下的

反例

设ξ、η是两个相互独立相同分布的随机变量

$$\xi \sim f(x)=\begin{cases}e^{-x}, & x\geqslant 0,\\ 0, & x<0.\end{cases} \qquad \eta \sim f(y)=\begin{cases}e^{-y}, & y\geqslant 0,\\ 0, & y<0.\end{cases}$$

令 $\quad \zeta=\xi\log(1+\eta)$

则 $\quad \mu_n=E\zeta^n=n!\gamma_n,$

其中 $\quad \gamma_n=\int_0^\infty [\log(1+x)]^n e^{-x}dx.$

可以证明，对$n\geqslant 1$有

$$e^{-1}\log(1+n)<\gamma_n^{\frac{1}{n}}<C\log(1+n). \qquad (2)$$

(见S. W. Dharmadhikeri; Am. Math Moth72(1965) P302—303)其中C是一个有限常数，因此

$$\left(\frac{\mu_n}{n!}\right)^{\frac{1}{n}}=\gamma_n^{\frac{1}{n}}>e^{-1}\log(1+n)\to\infty \quad (n\to\infty),$$

$\therefore$ 级数$\sum\limits_{n=1}^{\infty}\frac{\mu_n}{n!}s^n$不收敛（对任意非零$s$）

$\because \quad (n!)^{\frac{1}{n}}\sim e^{-1}n,$

由(2)

$\therefore \quad \mu_n^{\frac{1}{n}}\sim e^{-1}n\gamma_n^{\frac{1}{n}}<Ce^{-1}\log(1+n),$

于是 $\quad \sum\limits_{n=1}^{\infty}\mu_{2n}^{-\frac{1}{2n}}=\infty.$

再由Carleman条件（见J. A. shohat and T. D Tamarkin; The problem of Moments. Amer Math Soc. 1943 P19）知序列$\{\mu_n\}$唯一决定z之分布。

第四章　与数字特征有关的反例

4—1　随机变量的数学期望不存在（从而方差也不存在）的例子

有关概念

1　如果随机变量ξ的分布函数为$F(x)$，且

$$\int_{-\infty}^{\infty}|x|\,dF(x)<\infty,$$

则ξ的数学期望定义为

$$E\xi=\int_{-\infty}^{\infty}x\,dF(x).$$

等号右方是勒贝格——司梯阶积分。

（1）当ξ是离散型随机变量，且其密度矩阵为$\begin{pmatrix}x_1, & x_2, & x_3, & \cdots \\ p_1, & p_2, & p_3, & \cdots\end{pmatrix}$时，上述积分变为$\sum\limits_i x_ip_i$；

（2）当ξ是密度函数为$f(x)$的连续型随机变量时，上述积分变为$\int_{-\infty}^{\infty}xf(x)dx$。

2　若

$$\int_{-\infty}^{\infty}(x-E)\xi^2dF(x)<\infty,$$

记

$$D\xi=\int_{-\infty}^{\infty}(x-E)\xi^2dF(x),$$

称$D\xi$为ξ的方差。

当ξ分别是连续型或离散型随机变量时，上式分别化为

$$D\xi=\int_{-\infty}^{\infty}(x-\in\xi)^2 f(x)dx$$

和 $$D\xi=\sum_i (x_i-E\xi)p_i^2.$$

3 $D\xi=E\xi^2-(E\xi)^2$.

定理 设随机变量ξ的t阶矩存在，则随机变量ξ的s阶矩必存在（其中$0<s<t$）

证明 设ξ的分布函数为$F(x)$，由假设知

$$E|\xi|^t<\infty,$$

$$\therefore \quad E|\xi|^s=\int_{|x|^s\leq 1}|x|^s dF(x)+\int_{|x|^s>1}|x|^s dF(x)$$

$$\leqslant p\{|\xi|^s\leqslant 1\}+\int_{|x|^s>1}|x|^t dF(x)$$

$$\leqslant 1+E|\xi|^t<\infty.$$

定理表明：若随机变量的高阶矩存在，则它的低阶矩必存在。然而，定理的逆不真。

下面我们来看两个反例。一个是随相变量的数学期望不存在，从而方差也不存在的例子；一个是虽然随机变量的数学期望存在，但其方差不存在的例子。

反例

例1（哥西分布）设ξ的密度函数为

$$f(x)=\frac{1}{\pi}\frac{1}{(1+x^2)} \qquad (x\in R_1),$$

由于 $$\int_{-A}^{A}|x|f(x)dx=\int_{-A}^{A}|x|\frac{1}{\pi}\cdot\frac{1}{1+x^2}dx$$

$$=\frac{2}{\pi}\int_0^A\frac{x}{1+x^2}dx=\frac{2}{\pi}\cdot\frac{1}{2}\int_0^A\frac{d(1+x^2)}{1+x^2}dx$$

$$= \frac{1}{\pi}\ln(1+x^2)\Big|_0^A \to \infty \qquad (A\to\infty),$$

故$E\xi$不存在，从而$D\xi$也不存在。

例2 设离散型随机变量ξ的分布列由下式给出

$$p_i = p\left(\zeta = (-1)^{i+1}\cdot\frac{3^i}{i}\right) = \frac{2}{3^i} \qquad (i=1,2,3,\cdots),$$

虽然级数

$$\sum_i x_i p_i = \sum_{i=1}^{\infty}(-1)^{i+1}\cdot\frac{3^i}{i}\cdot\frac{2}{3^i} = \sum_{i=1}^{\infty}(-1)^{i+1}\cdot\frac{2}{i}$$

收敛，但是

$$\sum_{i=1}^{\infty}|x_i|p_i = \sum_{i=1}^{\infty}\left|(-1)^{i+1}\frac{3^i}{i}\right|\cdot\frac{2}{3^i}$$

$$= \sum_{i=1}^{\infty}\frac{2}{i} = \infty,$$

所以$E\xi$不存在，从而$D\xi$也不存在。

4—2 随机变量的数学期望存在，但方差不存在的例子

反例（$n=2$的t—分布）

设ξ的密度函数

$$f(x) = \frac{\Gamma(3/2)}{\sqrt{2\pi}}\left(1+\frac{x^2}{2}\right)^{-3/2}$$

因为 $$\int_{-\infty}^{\infty}|x|f(x)dx = \int_{-\infty}^{\infty}\frac{\Gamma(3/2)}{\sqrt{2\pi}}|x|(1+x^2/2)^{-3/2}dx$$

$$= \frac{\sqrt{2}\cdot\Gamma(3/2)}{\sqrt{\pi}}\int_0^{\infty}x\cdot\frac{1}{(1+x^2/2)^{3/2}}dx<\infty,$$

所以$E\xi$存在，且

$$E\xi=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{\infty}\frac{\Gamma\left(\frac{3}{2}\right)}{\sqrt{2\pi}}x\cdot\frac{1}{\left(1+\frac{x^2}{2}\right)^{3/2}}dx$$

$$=0,$$

但是

$$\int_{-\infty}^{\infty}(x-E\xi)^2 f(x)dx$$

$$=\int_{-\infty}^{\infty}\frac{\Gamma(3/2)}{\sqrt{2\pi}}\cdot x^2\cdot\frac{1}{\left(1+\frac{x^2}{2}\right)^{3/2}}dx=+\infty.$$

所以方差$D\xi$不存在。

进一步的讨论

对于一般的t—变量，若

$$\xi\sim t(n)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}\qquad n>1.$$

则当$r<n$时，$E\xi^r$存在，且

$$E\xi^r=\begin{cases}0, & \text{当}r\text{是奇数时，}\\ n^{\frac{r}{2}}\dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}, & \text{当}r\text{是偶数时。}\end{cases}$$

证明 当r为奇数时，$E\xi^r=0$为显然。当r为偶数时

$$E\xi^r=\int_{-\infty}^{\infty}x^r\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}dx$$

$$=2\int_0^{\infty} x^r \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx.$$

令 $$y=\frac{x^2}{n}\left(1+\frac{x^2}{n}\right)^{-1}=\frac{x^2}{n\left(1+\frac{x^2}{n}\right)}=\frac{x^2}{n^2+x^2},$$

则 $$x^2=\frac{ny}{1-y}, \qquad x=\left(\frac{ny}{1-y}\right)^{\frac{1}{2}},$$

$$dx=\frac{n^{\frac{1}{2}}y^{-\frac{1}{2}}}{2(1-y)^{\frac{3}{2}}}dy,$$

$$1+\frac{x^2}{n}=1+\frac{1}{n}\ \frac{ny}{1-y}$$

$$=1+\frac{y}{1-y}=\frac{1}{1-y}.$$

∴ $$E\xi^r=2\int_0^1\left(\frac{ny}{1-y}\right)^{\frac{r}{2}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma\left(\frac{n}{2}\right)}\left(\frac{1}{1-y}\right)^{-\frac{n+1}{2}}$$

$$\times\frac{n^{\frac{1}{2}}y^{-\frac{1}{2}}}{2(1-y)^{\frac{3}{2}}}dy$$

$$=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)}n^{\frac{r}{2}}\int_0^1 y^{\frac{r}{2}-\frac{1}{2}}(1-y)^{\frac{n+1}{2}-\frac{r}{2}-\frac{3}{2}}dy$$

$$=n^{\frac{r}{2}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^1 y^{\frac{1+r}{2}-1}(1-y)^{\frac{n-r}{2}-1}dy$$

$$= n^{\frac{r}{2}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \mathrm{B}\left(\frac{r+1}{2}, \frac{n-r}{2}\right) \text{（当} n > r \text{时）}$$

$$= n^{\frac{r}{2}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

（当$n > r$时）

$$= n^{\frac{r}{2}} \frac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \qquad \text{（当} n > r \text{时）。}$$

4—3 任何阶矩都不存在的随机变量的例子

有关概念

设$g(x)$是一元波雷尔可测函数，$F(x)$表随机变量ξ的分布函数，若

$$\int_{-\infty}^{\infty} |g(x)| dF(x) < \infty,$$

则$g(\xi)$的数学期望定义为

$$Eg(\xi) = \int_{-\infty}^{\infty} g(x) dF(x)。$$

对连续型的$F(x)$，若$f(x)$为其密度函数，则

$$Eg(\xi) = \int_{-\infty}^{\infty} g(x) f(x) dx。$$

对离散型的$F(x)$，若其密度矩阵为$\begin{pmatrix} a_1, & a_2, & \cdots \\ p_1, & p_2, & \cdots \end{pmatrix}$，

则 $$Eg(\xi)=\sum_i g(a_i)p_i.$$

将$g(x)$特殊化，就得到各种数字特征：

(1) 令$g(x)=x^k(k>0)$，称$E\xi^k$为ξ的k阶矩；

(2) 令$g(x)=|x|^k(k>0)$，称$E|\xi|^k$为ξ的k阶绝对矩。

下面举一个任何阶矩都不存在的随机变量的例子。

反例

考虑随机变量ξ，它的密度函数

$$f(x)=\begin{cases}\dfrac{1}{2|x|(\log|x|)^2}, & |x|>e,\\ 0, & 其他。\end{cases}$$

则可证对于任何$a>0$，$E|\xi|^a=\infty$。

事实上，对$\forall a>0$，由于

$$\lim_{x\to\infty}\frac{x^a}{(\log x)^2}=\infty.$$

所以$\exists M>e$，使当$x>M$时

$$\frac{x^a}{(\log x)^2}>1.$$

此时
$$\begin{aligned}E|\xi|^a&=\int_{-\infty}^{\infty}|x|^a f(x)dx\\&=2\int_e^{+\infty}\frac{x^a}{2x(\log x)^2}dx\\&\geqslant\int_M^{+\infty}\frac{x^a}{x(\log x)^2}dx\\&\geqslant\int_M^{\infty}\frac{1}{x}dx=\infty,\end{aligned}$$

所以 $E\xi^a$不存在（$a>0$）。

进一步的讨论

我们有下述结果：

定理 设ξ是概率空间$(\Omega, \mathscr{F}, P)$上的一个随机变量．设对某个$k>0$，$E|\xi|^k<\infty$，则

$$n^k \cdot P(|\xi|>n) \to 0 \qquad (n\to\infty)$$

(见〔2〕P.83)。

(1) 利用这一结果，对本款例子中的随机变量ξ可以证得$E|\xi|^\alpha=\infty$，对任意$\alpha>0$都成立．

事实上，ξ的分布函数为

$$F(x)=\begin{cases}\dfrac{1}{2\log|x|}, & x\leqslant -e,\\ \dfrac{1}{2}, & -e<x<e,\\ 1-\dfrac{1}{2\log x}, & x\geqslant e.\end{cases}$$

于是对$n>e$

$$P(|\xi|>n)=1-F(n)+F(-n)=\frac{1}{\log n},$$

所以对任何$\alpha>0$。

$$n^\alpha \cdot P(|\xi|>n)=n^\alpha\cdot\frac{1}{\log n}\to\infty \qquad (n\to\infty).$$

由此推出对每一个$\alpha>0$，$E|\xi|^\alpha=\infty$.

(2) 上述定理的逆命题一般不成立．也就是说，如果当$n\to\infty$时，对某个$k>0$，有

$$n^k P(|\xi|>n)\to 0$$

未必能推出$E|\xi|^k<\infty$.

例如，考虑随机变量ξ

$$P(\xi=n)=\frac{c}{n^2\log n}\quad(n=2,3,\cdots),$$

其中c是由

$$\sum_{n=2}^{\infty}\frac{c}{n^2\log n}=1$$

确定的常数。我们有

$$P\{\xi>n\}\approx c\int_n^{\infty}\frac{1}{x^2\log x}dx\approx c\cdot n^{-1}(\log n)^{-1}$$

（此处≈表示两边之比当$n\to\infty$时趋于1）且$nP(\xi>n)\to 0$，但是

$$E\xi=\sum_{n=3}^{\infty}\frac{c}{n\cdot\log n}=\infty.$$

(3) 但若$n^pP(|\xi|>n)\to 0$ $(n\to\infty)$ 对某个$p>0$成立，则可证明对于$0<q<p$，有

$$E|\xi|^q<\infty$$

（见〔1〕P38推论3）。

4—4 随机变量ξ的一阶矩存在，但没有更高整数阶的矩的例子

反例

设ξ的密度函数为

$$f(x)=\begin{cases}\dfrac{2}{x^3}, & x\geq 1,\\ 0, & x<1.\end{cases}$$

则 $$E\xi=\int_{-\infty}^{\infty}xf(x)dx=\int_1^{\infty}\frac{2}{x^2}dx=2.$$

但对任意整数$k>1$

$$E|\xi|^k=\int_{-\infty}^{\infty}|x|^k f(x)dx$$

$$=\int_1^{\infty}x^k\cdot\frac{2}{x^3}dx=\infty.$$

所以ξ没有比1更高的整数阶矩。

4—5 随机变量ξ在$0<r<1$时，$E\xi^r$ 存在；但在$r\geqslant1$时，$E\xi^r$ 不存在的例子

反例

设ξ服从哥西分布，其密度函数

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}\qquad(x\in R_1)$$

$$E|\xi|^r=\int_{-\infty}^{\infty}|x|^r f(x)dx$$

$$=\frac{2}{\pi}\int_0^{\infty}x^r\cdot\frac{1}{1+x^2}dx.$$

在积分中，设

$$t=\frac{1}{1+x^2},\qquad x=\left(\frac{1-t}{t}\right)^{\frac{1}{2}},$$

$$dt=-\frac{2x}{(1+x^2)^2}dx,\quad -\frac{2x}{1+x^2}dx=\frac{1}{t}dt,$$

所以 $$E|\xi|^r=-\frac{2}{2\pi}\int_0^{\infty}x^{r-1}\cdot\left(-\frac{2x}{1+x^2}\right)dx$$

$$=\frac{1}{\pi}\int_0^1\left(\frac{1-t}{t}\right)^{\frac{r-1}{2}}\cdot t^{-1}dt$$

$$=\frac{1}{\pi}\int_0^1 t^{\frac{1-r}{2}-1}(1-t)^{\frac{1+r}{2}-1}dt.$$

右端积分在$r<1$时收敛；在$r\geqslant 1$时发散即当$r<1$时，$E\xi^r$存在；$r\geqslant 1$时，$E\xi^r$不存在。

4—6 随机变量ξ_1和ξ_2，它们的一切整数阶矩都相同（即$E\xi_1^k=E\xi_2^k$，$k=1,2,3\cdots$）但它们的分布函数不相等的例子

有关概念和命题

1 对于正态分布、poisson 分布、二项分布等情况，它们的分布函数完全被一、两个参数所确定，而这些参数又都被随机变量的头两个矩所完全决定。因此，对这些分布，只要分布类型已知，那么由随机变量的一阶矩和二阶矩就可以完全决定其分布函数。如果我们不知道分布函数属于什么类型，则一般说来，不但知道头两阶矩还不能决定其分布函数，并且即使知道了一切整数阶矩仍不能决定。我们可以作出一切整数阶矩相同但分布函数不同的随机变量的例子。

2 对于连续型分布函数

$$F_1(x)=\int_{-\infty}^{x} f_1(y)dy \qquad F_2(x)=\int_{-\infty}^{x} f_2(y)dy$$

若$F_1(x)=F_2(x)$，$(x\in R_1)$，则几乎处处有（关于L—测度）$f_1(x)=f_2(x)$。

反例

设随机变量ξ_1和ξ_2分别有密度函数

$$f_1(x)=\begin{cases} c\{\exp[-x^{\mu}\cos\mu\pi]\}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

和
$$f_2(x)=\begin{cases} c[1+\sin(x^{\mu}\sin\mu\pi)]\exp[-x^{\mu}\cdot\cos\mu\pi], & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

其中　　$0<\mu<1/2$，$c=\dfrac{\mu(\cos\mu\pi)^{1/\mu}}{\Gamma(1/\mu)}$，

则　　$E\xi_1^n=E\xi_2^n \qquad (n=0,1,2,\cdots)$

但是，ξ_1和ξ_2分布函数$F_{\xi_1}(x)$和$F_{\xi_2}(x)$不相等。

我们分以下几步论述：

(1) 引理　假定$\mu>0$，$0<\alpha<\pi/2$，则对$n=0,1,2\cdots$

$$\int_0^\infty e^{-x^\mu\cos\alpha}\sin(x^\mu\sin\alpha)\cdot x^n dx$$

$$=\frac{1}{\mu}\Gamma\left(\frac{n+1}{\mu}\right)\sin\frac{(n+1)\alpha}{\mu}.$$

特别当$\alpha=\mu\pi$，$0<\mu<1/2$时，就得到

$$\int_0^\infty x^n\sin(x^\mu\sin\mu\pi)\exp(-x^\mu\cos\mu\pi)dx=0$$

$$(n=0,1,2,\cdots).$$

证明　记

$$I(\alpha)=\int_0^\infty e^{-x^\mu}e^{i\alpha}\cdot x^n dx$$

作变换　$z=x^\mu\cdot e^{i\alpha}$，$x^\mu=ze^{-i\alpha}$

$$x=z^{1/\mu}\cdot e^{-\alpha i/\mu},\quad dx=\frac{1}{\mu}\cdot z^{\frac{1}{\mu}-1}\cdot e^{-\frac{\alpha}{\mu}i}\cdot dz$$

所以　$$I(\alpha)=\int_L e^{-z}\cdot z^{\frac{n}{\mu}}\cdot e^{-\frac{\alpha n}{\mu}i}\cdot\frac{1}{\mu}\cdot z^{\frac{1}{\mu}-1}\cdot e^{-\frac{\alpha}{\mu}i}dz$$

$$=\frac{1}{\mu}e^{-\frac{(n+1)\alpha}{\mu}i}\int_L e^{-z}\cdot z^{\frac{n+1}{\mu}-1}dz.$$

这里L表示射线：$\arg z=\alpha$

因为对$z=re^{i\theta}$，$0\leqslant\theta\leqslant\alpha$来说，被积函数$e^{-z}z^{\frac{n+1}{\mu}-1}$，当$\frac{1}{r}$超向于0时，收敛于0，且收敛性关于$\theta$是一致的，所以得到

$$\int_L e^{-z}\cdot z^{\frac{n+1}{\mu}-1}dz$$

$$=\int_0^\infty e^{-x}\cdot x^{\frac{n+1}{\mu}-1}dx=\Gamma\left(\frac{n+1}{\mu}\right),$$

于是 $$\int_0^\infty e^{-x^\mu e^{i\alpha}}\cdot x^n dx=\frac{1}{\mu}e^{-i\frac{n+1}{\mu}\alpha}\Gamma\left(\frac{n+1}{\mu}\right),$$

即 $$\int_0^\infty e^{-x^\mu\cos\alpha}[\cos(x^\mu\sin\alpha)-i\sin(x^\mu\sin\alpha)]x^n dx$$

$$=\frac{1}{\mu}\left[\cos\frac{(n+1)\alpha}{\mu}-i\sin\frac{(n+1)\alpha}{\mu}\right]\Gamma\left(\frac{n+1}{\mu}\right).$$

从而得到

$$\int_0^\infty e^{-x^\mu\cos\alpha}\cdot\sin(x^\mu\sin\alpha)x^n dx$$

$$=\frac{1}{\mu}\sin\left(\frac{(n+1)\alpha}{\mu}\right)\Gamma\left(\frac{n+1}{\mu}\right)\quad(n=0,1,2,3,\cdots).$$

当$0<\mu<\frac{1}{2}$时，若取$\alpha=\mu\pi$，就有

$$\int_0^\infty x^n\cdot\sin(x^\mu\sin\mu\pi)\cdot\exp(-x^\mu\cos\mu\pi)dx=0$$

$$(n=0,1,2,3,\cdots).$$

(2) $E\xi_1^n=E\xi_2^n\quad(n=0,1,2,3,\cdots)$.

证明 $$E\xi_1^n=\int_0^\infty x^n f_1(x)dx$$

$$=\int_0^\infty x^n\cdot C\cdot\exp(-x^\mu\cos\mu\pi)dx$$

(令 $x^\mu\cos\mu\pi=t$，$x=\left(\frac{t}{\cos\mu\pi}\right)^{\frac{1}{\mu}}$)

$$=\int_0^\infty C\cdot\left(\frac{t}{\cos\mu\pi}\right)^{n/\mu}\cdot e^{-t}\cdot\frac{1}{\mu}(\cos\mu\pi)^{-\frac{1}{\mu}}\cdot t^{\frac{1}{\mu}-1}dt$$

$$=C\cdot\frac{1}{\mu}(\cos\mu\pi)^{-\frac{n+1}{\mu}}\int_0^\infty t^{\frac{n+1}{\mu}-1}\cdot e^{-t}dt$$

$$= \frac{C}{\mu}(\cos\mu\pi)^{-\frac{n+1}{\mu}}\Gamma\left(\frac{n+1}{\mu}\right)$$

$$= \frac{\mu(\cos\mu\pi)^{1/\mu}}{\Gamma\left(\frac{1}{\mu}\right)}\cdot\frac{1}{\mu}(\cos\mu\pi)^{-\frac{n+1}{\mu}}\Gamma\left(\frac{n+1}{\mu}\right)$$

$$= \frac{\Gamma\left(\frac{n+1}{\mu}\right)}{\Gamma\left(\frac{1}{\mu}\right)}(\cos\mu\pi)^{-\frac{n}{\mu}} \quad (n=0,1,2,\cdots).$$

由引理得：

$$E\xi_2^n = \int_0^\infty x^n f_2(x)dx$$

$$= \int_0^\infty x^n\cdot c[1+\sin(x^\mu\cdot\sin\mu\pi)]\exp(-x^\mu\cos\mu\pi)dx$$

$$= \int_0^\infty x^n\cdot c\cdot\exp(-x^\mu\cos\mu\pi)dx$$

$$+\int_0^\infty x^n\cdot c\cdot\sin(x^\mu\cdot\sin\mu\pi)\exp(-x^\mu\cos\mu\pi)dx$$

$$= \int_0^\infty cx^n\exp(-x^\mu\cos\mu\pi)dx$$

$$= E\xi_1^n \qquad (n=0,\ 1,\ 2,3\cdots).$$

(3) $F_{\xi_1}(x) \neq F_{\xi_2}(x)$.

证明 若$F_{\xi_1}(x) = F_{\xi_2}(x)$，则几乎处处有$f_1(x) = f_2(x)$，即

$$m\{x|f_1(x)\neq f_2(x)\}=0, \qquad (*)$$

其中m为R_1上的勒贝格测度.

但是 $\{x|f_1(x)\neq f_2(x)\}$

$$=\{x|x>0,\ \sin(x^\mu\sin\mu\pi)\neq 0\}$$

$$=\{x|x>0,\ x^\mu\cdot\sin\mu\pi\neq k\pi,\ k=0,\pm1,\pm2,\ \cdots\}$$

$$=\left\{x|x>0,\ x^\mu\neq\frac{k\pi}{\sin(\mu\pi)},\ k=0,\pm1,\ \pm2,\cdots\right\}$$

$$=\left\{x\neq\left(\frac{k\pi}{\sin\mu\pi}\right)^{1/2\mu},\ x>0,\ k=0,\pm1,\pm2,\cdots\right\}$$

由于 $\left\{x=\left(\frac{k\pi}{\sin\mu\pi}\right)^{1/\mu},\ x>0,\ k=0,\pm1,\pm2,\cdots\right\}$

是可列集，故

$$m\left\{x=\left(\frac{k\pi}{\sin\mu\pi}\right)^{1/\mu},\ x>0,\ k=0,\pm1,\pm2\cdots\right\}=0.$$

从而 $m\{x|f_1(x)\neq f_2(x)\}\neq0$。

与(＊) 相矛盾．因此得到

$$F_{\xi_1}(x)\neq F_{\xi_2}(x)\qquad\qquad(x\in R_1).$$

进一步的讨论

在某些条件下，由一切整数阶矩可以唯一确定分布函数

(1°)(见例*3-25之定理)

(2°)如果随机变量ξ以概率1有界，即存在常数$c>0$，使$p(|\xi|<c)=1$，则ξ的分布函数$F(x)$被它的各阶矩$\{a_n\}$唯一决定。事实上，这时 $F(-c)=p(\xi<-c)=0$，$F(c)=1$，

故 $$|a_n|\leqslant\int_{-c}^{c}|x|^n dF(x)\leqslant c^n\int_{-c}^{c}dF(x)=c^n.$$

从而 $$\sum_{n=0}^{\infty}\left|\frac{a_nc^n}{n!}\right|\leqslant\sum_{n=0}^{\infty}\frac{c^{2n}}{n!}=e^{c^2}<\infty.$$

4—7 随机变量ξ_1和ξ_2不相关，但也不独立的例子

有关概念和命题

1 定义 设随机变量ξ_1和ξ_2的方差$D\xi_1$和$D\xi_2$均存在且大于0，记

$$\rho(\xi_1,\xi_2)=\frac{E[(\xi_1-E\xi_1)(\xi_2-E\xi_2)]}{\sqrt{D\xi_1}\cdot\sqrt{D\xi_2}}$$

称$\rho(\xi_1,\xi_2)$为ξ_1和ξ_2的相关系数。

数字特征$\rho(\xi_1,\xi_2)$描述了随机变量ξ_1和ξ_2的线性相关程度。当$\rho(\xi_1,\xi_2)=0$时，称ξ_1和ξ_2不相关。

2 定理 若ξ_1、ξ_2独立，则ξ_1和ξ_2不相关.但反过来则不一定正确。

反例 设(ξ_1,ξ_2)均匀分布于以坐标原点为中心，r_0为半径的圆的内部，则$\rho(\xi_1,\xi_2)=0$，但ξ_1和ξ_2不独立。

证明如下：因(ξ_1,ξ_2)为均匀分布，故(ξ_1,ξ_2)的密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi r_0^2}, & \text{当}x^2+y^2\leqslant r_0^2,\\ 0, & \text{当}x^2+y^2>r_0^2.\end{cases}$$

注意到边际分布密度

$$f_{\xi_1}(x)=\int_{-\sqrt{r_0^2-x^2}}^{\sqrt{r_0^2-x^2}}\frac{1}{\pi r_0^2}dy=\frac{2\sqrt{r_0^2-x^2}}{\pi r_0^2},(-r_0\leqslant x\leqslant r_0)$$

$$f_{\xi_2}(y)=\int_{-\sqrt{r_0^2-y^2}}^{\sqrt{r_0^2-y^2}}\frac{1}{\pi r_0^2}dy=\frac{2\sqrt{r_0^2-y^2}}{\pi r_0^2},(-r_0\leqslant y\leqslant r_0).$$

显然 $f(x,y)\neq f_{\xi_1}(x)\cdot f_{\xi_2}(y)$，

故ξ_1，ξ_2不独立。但

$$E\xi_1=\int_{-\infty}^{\infty}xf_{\xi_1}(x)dx=\int_{-\infty}^{\infty}\frac{2x\sqrt{r_0^2-x^2}}{\pi r_0^2}dx=0.$$

同理可得 $E\xi_2=0$.

又由于$f(x,y)$的对称性，可知

$$E[(\xi_1-E\xi_1)(\xi_2-E\xi_2)]$$

$$=\iint\limits_{x^2+y^2\leqslant r_0^2}xy\frac{1}{\pi r_0^2}dxdy=0,$$

从而ξ_1和ξ_2的相关系数$\rho(\xi_1, \xi_2)=0$。

进一步的讨论

在一些特殊的场合下，独立性与不相关性是等价的：

(1) 若(ξ_1, ξ_2)服从二元正态分布，则ξ_1，ξ_2独立$\Longleftrightarrow$ $\rho(\xi_1,\xi_2)=0$。

(2) 若ξ_1，ξ_2都只取两个值，则不相关必独立。

证明 设

$$p(\xi_1=x_1)=p_1,\ p(\xi_1=x_2)=1-p_1,$$
$$p(\xi_2=y_1)=p_2,\ p(\xi_2=y_2)=1-p_2.$$

其中 $x_1 \neq x_2$，$y_1 \neq y_2$。

则(ξ_1,ξ_2)的联合概率分布和ξ_1、ξ_2的边缘概率分布可由下表给出

ξ_2 \ ξ_1	x_1	x_2	ξ_2的边缘分布
y_1	$p(\xi_1=x_1,\xi_2=y_1)$	$p(\xi_1=x_2,\xi_2=y_1)$	p_2
y_2	$p(\xi_1=x_1,\xi_2=y_2)$	$p(\xi_1=x_2,\xi_2=y_2)$	$1-p_2$
ξ_1的边缘分布	p_1	$1-p_1$	

由上表可知

$$(I)\begin{cases} p(\xi_1=x_1,\xi_2=y_1)+p(\xi_1=x_1,\xi_2=y_2)=p_1, & (1)\\ p(\xi_1=x_2,\xi_2=y_1)+p(\xi_1=x_2,\xi_2=y_2) & \\ =1-p_1, & (2)\\ p(\xi_1=x_1,\ \xi_2=y_1)+p(\xi_1=x_2,y_2=y_1)=p_2, & (3)\\ p(\xi_1=x_1,\xi_2=y_2)+p(\xi_1=x_2,\ \xi_2=y_2) & \\ =1-p_2.(4) & \end{cases}$$

将$p(\xi_1=x_i,\ \xi_2=y_j)(i,\ j=1,2)$看作未知数时，$(I)$中有三个

方程是独立的

$\because \quad \rho(\xi_1,\ \xi_2)=0,$

$\therefore \quad \operatorname{cov}(\xi_1,\xi_2)=0.$

即 $$\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i-E\xi_1)(y_j-E\xi_2)p(\xi_1=x_i,\xi_2=y_j)=0,$$

也即
$$\begin{aligned}&(x_1-E\xi_1)(y_1-E\xi_2)p(\xi_1=x_1,\xi_2=y_1)\\&+(x_1-E\xi_1)(y_2-E\xi_2)p(\xi_1=x_1,\ \xi_2=y_2)\\&+(x_2-E\xi_1)(y_1-E\xi_2)p(\xi_1=x_2,\xi_2=y_1)\\&+(x_2-E\xi_1)(y_2-E\xi_2)p(\xi_1=x_2,\ \xi_2=y_2)\\=&\ 0.\end{aligned}\tag{5}$$

其中 $E\xi_1=x_1p_1+x_2(1-p_1),$

$E\xi_2=y_1p_2+y_2(1-p_2).$

将(1)、(2)、(3)、(5)联立得：

$$(II)\quad\begin{cases}p(\xi_1=x_1,\xi_2=y_1)+p(\xi_1=x_1,\xi_2=y_2)=p_1,\\p(\xi_1=x_2,\xi_2=y_1)+p(\xi_1=x_2,\xi_2=y_2)=1-p_1,\\p(\xi_1=x_1,\xi_2=y_1)+p(\xi_1=x_2,\xi_2=y_1)=p_2.\\\qquad(5)\end{cases}$$

方程组(II)的系数矩阵

$$\begin{pmatrix}1 & 1\\0 & 0\\1 & 0\\(x_1-E\xi_1)(y_1-E\xi_2) & (x_1-E\xi_1)(y_2-E\xi_2)\\0 & 0\\1 & 1\\1 & 1\\(x_2-E\xi_1)(y_1-E\xi_2) & (x_2-E\xi_1)(y_2-E\xi_2)\end{pmatrix}$$

是满秩的，故方程組(II)有唯一解。但

$$
(III)\begin{cases}
p(\xi_1=x_1,\xi_2=y_1)=p_1p_2,\\
p(\xi_1=x_1,\xi_2=y_2)=p_1(1-p_2),\\
p(\xi_1=x_2,\xi_2=y_1)=(1-p_1)p_2,\\
p(\xi_1=x_2,\xi_2=y_2)=(1-p_1)(1-p_2)
\end{cases}
$$

也满足(II)，故(III)就是(II)的唯一解。

(III) 说明了ξ_1、ξ_2是相互独立的，证毕。

4—8 相关系数$\rho(\xi_1,\xi_2)>0$，$\rho(\xi_2,\xi_3)>0$，但$\rho(\xi_1,\xi_3)<0$的例子

有关概念

1 相关系数概念见4—7。

2 n维正态分布．设已给n阶对称正定矩阵$B=(b_{ZK})$，又给向量$a=(a_1,\ \cdots,\ a_n)$，a_i为任意实数$(i=1,2,\cdots,n)$，定义n元函数：

$$
f(x)=\frac{1}{(2\pi)^{n/2}|B|^{1/2}}\exp\left\{-\frac{1}{2}(x-a)B^{-1}(x-a)^{\mathrm{T}}\right\}.\ (*)
$$

其中向量$x=(x_1,x_2,\cdots,x_n)$，B^{-1}表B的逆矩阵，$|B|$表B的行列式的值．称以$(*)$中的$f(x)$为密度函数的连续型分布为n维正态分布。

3 如果随机向量$(\xi_1,\xi_2,\cdots,\xi_n)$服从以$(*)$中的$f(x)$为密度函数的$n$维正态分布，则

$$
\begin{aligned}
&a_i=E\xi i,\\
&b_{ij}=\mathrm{cov}(\xi_i,\xi_j)\triangleq E[(\xi_i-E\xi_i)(\xi_j-E\xi_j)].\\
&\qquad(i,j=1,2,\cdots,n)
\end{aligned}
$$

并称n阶矩阵$B=(b_{ij})$为$(\xi_1,\xi_2,\cdots,\xi_n)$的协方差矩阵.

众所周知，相关系数是描述随机变量之间线性相关程度的一种数字特征。若有三个随机变量ξ_1、ξ_2和ξ_3，当$\rho(\xi_1,\xi_2)>0$，$\rho(\xi_2,\xi_3)>0$时，人们一般料想应有$\rho(\xi_1,\xi_3)>0$。但是事实并非如此．请看下面的反例：

反例

设随机向量(ξ_1,ξ_2,ξ_3)服从三维正态分布，协方差矩阵为

$$B=\begin{pmatrix}1 & 0.1 & -0.1\\ 0.1 & 1 & 0.1\\ -0.1 & 0.1 & 1\end{pmatrix}$$

显然B为三阶对称阵．同时也不难验证B是正定矩阵。由于$D\xi_1=D\xi_2=D\xi_3=1$，所以

$$\rho(\xi_1,\xi_2)=b_{12}=0.1>0,$$

$$\rho(\xi_2,\xi_3)=b_{23}=0.1>0.$$

但是却有$\rho(\xi_1,\xi_3)=b_{13}=-0.1<0$。

4—9 ξ_1和ξ_2不独立，但$E(\xi_1\cdot\xi_2)=E\xi_1\cdot E\xi_2$的例子

有关概念

1 随机变量ξ_1、ξ_2独立的充要条件是对任意一维波雷尔点集B_1、B_2，有

$$p(\xi_1\in B_1,\ \xi_1\in B_2)=p(\xi_1\in B_1)\cdot p(\xi_2\in B_2)$$

成立。

2 若ξ_1、ξ_2独立，则

$$E(\xi_1\xi_2)=E\xi_1\cdot E\xi_2,$$

但是ξ_1、ξ_2独立仅是上式成立的充分条件，并非必要条件。我们找到ξ_1、ξ_2不独立但

$$E(\xi_1\xi_2)=E\xi_1\cdot E\xi_2$$

的例子。

反例

取$\Omega=[0.1]$，$\mathscr{F}$为$[0,1]$中的波雷尔点集类，p为直线上的勒贝格测度。

定义$\xi(\omega)=\omega$，则$\xi(\omega)$是$(\Omega,\mathscr{F},P)$上的随机变量。又对任意$0\leqslant x\leqslant 1$，

$$F_\xi(x)=p(\xi(\omega)<x)=p(0\leqslant\omega<x)=x.$$

因此$\xi(\omega)$服从$[0,1]$上的均匀分布。

再作两个随机变量：

$$\xi_1(\omega)=g_1(\xi(\omega))=\sin 2\pi\omega,$$

$$\xi_2(\omega)=g_2(\xi(\omega))=\cos 2\pi\omega,$$

于是

$$E\xi_1=\int_{-\infty}^{\infty}g_1(x)dF_\xi(x)$$

$$=\int_0^1\sin 2\pi x\cdot 1dx=0,$$

$$E\xi_2=\int_0^1\cos 2\pi x\cdot 1dx=0,$$

$$E\xi_1\xi_2=\int_{-\infty}^{\infty}g_1(x)\cdot g_2(x)dF_\xi(x)$$

$$=\frac{1}{2}\int_0^1\sin 4\pi x dx=0.$$

故有 $E\xi_1\xi_2=E\xi_1\cdot E\xi_2$.

取ε足够小，使事件

$$A_{\xi_1}=\{\omega\,|\,|\xi_1(\omega)-1|<\varepsilon\}$$

$$= \{\omega \mid 1-\varepsilon < \sin 2\pi\omega < 1+\varepsilon\}$$

和 $$B_{\xi_2} = \{\omega \mid |\xi_2(\omega)+1| < \varepsilon\}$$

$$= \{\omega \mid -1-\varepsilon < \cos 2\pi\omega < -1+\varepsilon\}$$

互斥，即$A_{\xi_1} \cap B_{\xi_2} = \phi$，

从而 $$p(A_{\xi_1} \cap B_{\xi_2}) = 0.$$

而另一方面

$$p(A_{\xi_1}) = p(\omega \mid 1-\varepsilon < \sin 2\pi\omega < 1+\varepsilon) \neq 0,$$

$$p(B_{\xi_2}) = p(\omega \mid -1-\varepsilon < \cos 2\pi\omega < -1+\varepsilon) \neq 0,$$

于是 $$p(A_{\xi_1} \cap B_{\xi_2}) \neq p(A_{\xi_1}) \cdot p(B_{\xi_2}),$$

所以ξ_1和ξ_2不独立。

4—10 ξ_1、ξ_2独立，但$E(\xi_1+\xi_2)^k \neq E\xi_1^k + E\xi_2^k$的例子（正整数$k \neq 1$）

有关概念

对随机变量ξ_1、ξ_2有

$$E(\xi_1+\xi_2) = E\xi_1 + E\xi_2.$$

特别当ξ_1、ξ_2独立时

$$D(\xi_1+\xi_2) = D\xi_1 + D\xi_2.$$

但是由ξ_1、ξ_2独立，却推不出

$$E(\xi_1+\xi_2)^k = E\xi_1^k + E\xi_2^k \qquad (k \neq 1).$$

反例

设$\xi_1 \sim N(\mu_1, \sigma_1^2)$，　$\xi_2 \sim N(\mu_2, \sigma_2^2)$，其中$\mu_1 \neq 0$，$\mu_2 \neq 0$，且$\xi_1$、$\xi_2$相互独立。则

$$\xi_1+\xi_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2).$$

但是 $$E(\xi_1+\xi_2)^2 = D(\xi_1+\xi_2) + [E(\xi_1+\xi_2)]^2$$

$$=(\sigma_1^2+\sigma_2^2)+(\mu_1+\mu_2)^2,$$

可是 $E\xi_1^2=D\xi_1+(E\xi_1)^2=\sigma_1^2+\mu_1^2,$

$$E\xi_2^2=D\xi_2+(E\xi_2)^2=\sigma_2^2+\mu_2^2,$$

由于$\mu_1\neq0$，$\mu_2\neq0$，所以

$$E(\xi_1+\xi_2)^2\neq E\xi_1^2+E\xi_2^2.$$

进一步的讨论

上述反例表明，在独立随机变量相加时，和的矩一般不等于各项矩的和。对于独立随机变量ξ_1和ξ_2之和的矩应有以下等式：

$$E(\xi_1+\xi_2)^n=\sum_{k=0}^{n}C_n^k E\xi_1^k\cdot E\xi_2^{n-k}.$$

在随机变量的其他数字特征中，半不变量是一种重要的数字特征。随机变量的特征函数的对数在0点的k阶导数乘以i^{-k}叫做随和变量的k阶半不变量。

各阶的半不变量具有这样的性质：独立随机变量之和的半不变量等于同阶的各项半不变量之和。这是它比矩方便之处。

有关半不变量的性质的讨论以及它与矩之间的关系可参阅〔5〕p113—114。

4—11 $E[E(\eta|\xi)]$存在，但$E\eta$不存在，因而$E\eta=E[E(\eta|\xi)]$不成立的例子

有关概念和命题

1 两个随机变量ξ和η，假定它们具有联合密度函数$f(x,y)$，并以$f(y|x)$记已知在$\xi=x$的条件下，η的条件密度函数，以

$f_\xi(x)$记ξ的密度函数，则

$$f(y|x)=\frac{f(x,\ y)}{f_\xi(x)}.$$

2 在$\xi=x$的条件下，η的条件数学期望定义为

$$E[\eta|\xi=x]=\int_{-\infty}^{\infty}yf(y|x)dy.$$

3 若以$E(\eta|\xi)$记随机变量ξ的如下的函数：当$\xi=x$时，它取值为$E[\eta|\xi=x]$. 这样定义的$E(\eta|\xi)$是一个随机变量. 而且当$E\eta$存在时，有下列关系式

$$E\eta=E[E(\eta|\xi)].$$

这是条件数学期望的一个非常重要的性质. 但是，如果$E[E(\eta|\xi)]$存在而$E\eta$不存在时，则

$$E\eta=E[E(\eta|\xi)]$$

不再成立. 请看下面的反例.

反例

设ξ是一取正值的随机变量，其密度

$$f_\xi(x)=\begin{cases}\dfrac{\left(\frac{1}{2}\right)^{\frac{1}{2}}x^{-\frac{1}{2}}e^{-\frac{x}{2}}}{\Gamma\left(\frac{1}{2}\right)}, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

在给定$\xi=x(>0)$的条件下，η的条件密度函数为

$$f(y|x)=\frac{1}{\sqrt{2\pi}}x^{\frac{1}{2}}e^{-\frac{xy^2}{2}}\quad(-\infty<y<+\infty).$$

因为$f(y|x)$是y的对称函数，所以

$$E[\eta|\xi=x]=\int_{-\infty}^{\infty}yf(y|x)dy=0,$$

故$E[E(\eta|\xi)]$存在. 又由于η的边际密度

$$f_\eta(y)=\int_{-\infty}^{\infty} f_\xi(x)\cdot f(y|x)dx$$

$$=\int_0^{\infty}\frac{\left(\frac{1}{2}\right)^{\frac{1}{2}}x^{-\frac{1}{2}}e^{-\frac{x}{2}}}{\Gamma\left(\frac{1}{2}\right)}\cdot\frac{1}{\sqrt{2\pi}}x^{\frac{1}{2}}e^{-\frac{xy^2}{2}}dx$$

$$=\frac{1}{\pi(1+y^2)}\qquad(-\infty<y<+\infty).$$

因为 $\int_{-\infty}^{\infty}|y|\frac{1}{\pi(1+y^2)}dy=\infty$,

故$E\eta$不存在。所以$E\eta=E[E(\eta|\xi)]$不能成立。

4—12 中位数不唯一的例子

有关概念

对随机变量ξ，如果一个实数x满足下列条件 $p(\xi\leqslant x)\geqslant\frac{1}{2}$

和 $p(\xi\geqslant x)\geqslant\frac{1}{2}$,

则称x为随机变量的一个中位数，记为$x_{\frac{1}{2}}$.它是反映随机变量集中位置的一个数字特征。

中位数总存在，但中位数不唯一。

反例

离散型随机变量ξ的密度矩阵为

$$\begin{pmatrix}-2 & 0 & 1 & 2\\ \frac{1}{4} & \frac{1}{4} & \frac{1}{3} & \frac{1}{6}\end{pmatrix},$$

则 $$p(\xi\leqslant 0)=\frac{1}{2},\quad p(\xi\geqslant 0)=\frac{3}{4}>\frac{1}{2},$$

所以$x=0$是ξ的一个中位数。其实，对任何实数$0<x<1$，也都有

$$p(\xi\leqslant x)=p(\xi=-2)+p(\xi=0)=\frac{1}{2},$$

$$p(\xi\geqslant x)=p(\xi=1)+p(\xi=2)=\frac{1}{2},$$

所以对满足$0\leqslant x<1$的每一个实数x，它都是ξ的中位数。

4—13 数学期望不存在，但中位数存在的随机变量的例子

反例

设ξ服从哥西分布，密度函数为

$$f(x)=\frac{1}{\pi}\frac{1}{1+x^2}\quad(x\in R_1).$$

前已证明$E\xi$不存在，但是

$$p(\xi\geqslant 0)=\int_0^{+\infty}\frac{1}{\pi}\frac{1}{1+x^2}dx$$

$$=\frac{1}{\pi}\operatorname{arctg} x\Big|_0^{+\infty}=\frac{1}{2},$$

$$p(\xi\leqslant 0)=\int_{-\infty}^{0}\frac{1}{\pi}\frac{1}{1+x^2}dx=\frac{1}{2},$$

所以$x=0$是ξ的中位数。

4—14 随机变量的众数不唯一的例子

有关概念

如果随机变量ξ是离散型的，其分布列为

$$\begin{pmatrix} a_1, & a_2, & a_3, & \cdots \\ p_1, & p_2, & p_3, & \cdots \end{pmatrix}.$$

若$p_k=\sup(p_1, p_2, \cdots)$，则称$a_k$为$\xi$的众数（也称为$\xi$的最可能值）；如果$\xi$是连续型的，密度函数为$f(x)$，则称使$f(x)$达到最大值的点$x$为$\xi$的众数。

众数可以不唯一。

反例

例1 设ξ服从$[0, 1]$上的均匀分布，其密度函数为

$$f(x)=\begin{cases} 1, & x\in[0, 1], \\ 0, & \text{其他}. \end{cases}$$

则对任何$0\leqslant x\leqslant 1$，都是ξ的众数。

例2 设离散型随机变量ξ服从二项分布$B(8, \frac{1}{3})$，即有

$$p_m=p(\xi=m)=c_8^m\left(\frac{1}{3}\right)^m\left(\frac{2}{3}\right)^{8-m}$$

$$(m=0, 1, 2, \cdots, 8),$$

因为 $np-q=8\times\frac{1}{3}-\frac{2}{3}=2$

是整数，所以p_m在

$$k_0=[np-q]+1=3$$

和 $k_0-1=2$

处同时达到最大值。因此ξ的众数有两个，分别为2和3。

第五章　与特征函数、母函数有关的反例

5—1　随机变量ξ_1、ξ_2不独立，但$\varphi_{\xi_1+\xi_2}(t)=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t)$成立的例子

有关概念和命题

1　如果随机变量ξ的分布函数是$F(x)$，则称$e^{it\xi}$的数学期望

$$\varphi(t)\triangleq Ee^{i\xi t}$$

为ξ的特征函数。$\varphi(t)$是实变量t的复函数。当ξ为连续型随机变量时，

$$\varphi(t)=\int_{-\infty}^{\infty}e^{itx}f(x)dx;$$

当ξ为离散型随机变量时，

$$\varphi(t)=\sum_k e^{itx_k}\cdot p(\xi=x_k).$$

任一随机变量的特征函数总是存在的。

2　若$\eta=a\xi+b$，a、b是常数，则η的特征函数$\varphi_\eta(t)=e^{ibt}\varphi_\xi(at)$。

3　设$(\xi_1,\ \xi_2)$有联合密度函数$f(x,\ y)$，则$\xi=\xi_1+\xi_2$的密度函数为

$$f_{\xi_1+\xi_2}(z)=\int_{-\infty}^{\infty}f(x,\ z-x)dx.$$

4 若ξ_1、ξ_2独立，则

$$\varphi_{\xi_1+\xi_2}(t)=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t) \qquad (*)$$

值得注意的是上述条件(*)仅是ξ_1、ξ_2独立的必要条件，并非充分条件。下面的两个例子都表明，即便条件(*)成立，ξ_1和ξ_2仍可不独立。

反例

例1 设$(\xi_1,\ \xi_2)$有联合密度函数

$$f(x,\ y)=\begin{cases}\dfrac{1}{4}[1+xy(x^2-y^2)], & |x|\leqslant 1,\quad |y|\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

则ξ_1、ξ_2的密度$f_{\xi_1}(x)$、$f_{\xi_2}(y)$在$|x|\leqslant 1$及$|y|\leqslant 1$上分别为

$$f_{\xi_1}(x)=\int_{-1}^{1}\frac{1}{4}[1+xy(x^2-y^2)]dy=\frac{1}{2},$$

$$f_{\xi_2}(y)=\int_{-1}^{1}\frac{1}{4}[1+xy(x^2-y^2)]dx=\frac{1}{2},$$

故ξ_1、ξ_2都在$[-1,\ 1]$中服从均匀分布。既然

$$f_{\xi_1}(x)\cdot f_{\xi_2}(y)=\frac{1}{4}\neq f(x,\ y)\quad (|x|\leqslant 1,\quad |y|\leqslant 1),$$

所以ξ_1和ξ_2不独立。

由随机变量和的分布公式可知，$\xi_1+\xi_2$的密度函数为

$$f_{\xi_1+\xi_2}(z)=\int_{-\infty}^{\infty}f(x,\ z-x)dx.$$

由于$f(x,\ y)$在矩形$|x|\leqslant 1$，$|y|\leqslant 1$以外为0，故对每个固定的z，上式积分要在集合

$$E=\{x\,|\,|x|\leqslant 1,\ |x-z|\leqslant 1\}$$

上进行。

当 $-2\leqslant z\leqslant 0$ 时，

$$E=\{x\mid -1\leqslant x\leqslant z+1\}.$$

$0\leqslant z\leqslant 2$ 时，

$$E=\{x\mid z-1\leqslant x\leqslant 1\}.$$

$|z|>2$ 时，$E=\phi$。所以，当 $-2\leqslant z\leqslant 0$ 时，

$$f_{\xi_1+\xi_2}(z)=\int_{-1}^{z+1}\frac{1}{4}\left\{1+x(z-x)[x^2-(z-x)^2]\right\}dx$$

$$=\int_{-1}^{z+1}\frac{1}{4}[1+3x^2z^2-2x^3z-xz^3]dx$$

$$=\frac{1}{4}(2+z).$$

当 $0\leqslant z\leqslant 2$ 时，

$$f_{\xi_1+\xi_2}(z)=\int_{z-1}^{1}\frac{1}{4}[1+3x^2z^2-2x^3z-xz^3]dx$$

$$=\frac{1}{4}(2-z).$$

当 $|z|>2$ 时，

$$f_{\xi_1+\xi_2}(z)=0.$$

综上所述

$$f_{\xi_1+\xi_2}(z)=\begin{cases}\frac{1}{4}(2+z), & -2\leqslant z\leqslant 0,\\ \frac{1}{4}(2-z), & 0<z\leqslant 2,\\ 0, & |z|>2.\end{cases}$$

于是 $\xi_1+\xi_2$ 的特征函数

$$\varphi_{\xi_1+\xi_2}(t)=\int_{-\infty}^{\infty}e^{itz}f_{\xi_1+\xi_2}(z)dz$$

$$=\frac{1}{4}\left[\int_{-2}^{0}e^{itz}(2+z)dz+\int_{0}^{2}(2-z)e^{itz}dz\right]$$

$$=\frac{1}{4}\left[\frac{2-e^{2it}-e^{-2it}}{t^2}\right]$$

$$=\frac{1}{2t^2}(1-\cos 2t)$$

$$=\left(\frac{\sin t}{t}\right)^2.$$

而ξ_1和ξ_2的特征函数分别为

$$\varphi_{\xi_1}(t)=\int_{-1}^{1}\frac{1}{2}e^{itx}dx$$

$$=\frac{e^{itx}}{2it}\Big|_{-1}^{1}=\frac{1}{2it}(e^{it}-e^{-it})$$

$$=\frac{\sin t}{t},$$

$$\varphi_{\xi_2}(t)=\int_{-1}^{1}\frac{1}{2}e^{itx}dx=\frac{\sin t}{t},$$

于是得　$\varphi_{\xi_1+\xi_2}(t)=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t)$.

例2　设随机变量ξ服从$\lambda=1$, $\mu=0$的哥西分布，其密度函数为

$$f_\xi(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},$$

且$\eta=\xi$，则有

$$\varphi_{\xi+\eta}(t)=\varphi_\xi(t)\cdot\varphi_\eta(t),$$

但是ξ和η不独立.

事实上，ξ的特征函数

$$\varphi_\xi(t)=\int_{-\infty}^{\infty}e^{itx}\cdot\frac{1}{\pi}\ \frac{1}{1+x^2}dx. \tag{1}$$

当$t>0$时，考虑积分路径L为图 5—1 中箭头指示方向的复变量

积分

$$\int_L e^{itz}\frac{1}{1+z^2}dz. \quad (**)$$

若z位于上半圆周上，则$z=Re^{i\theta}$，$dz=iRe^{i\theta}d\theta$，

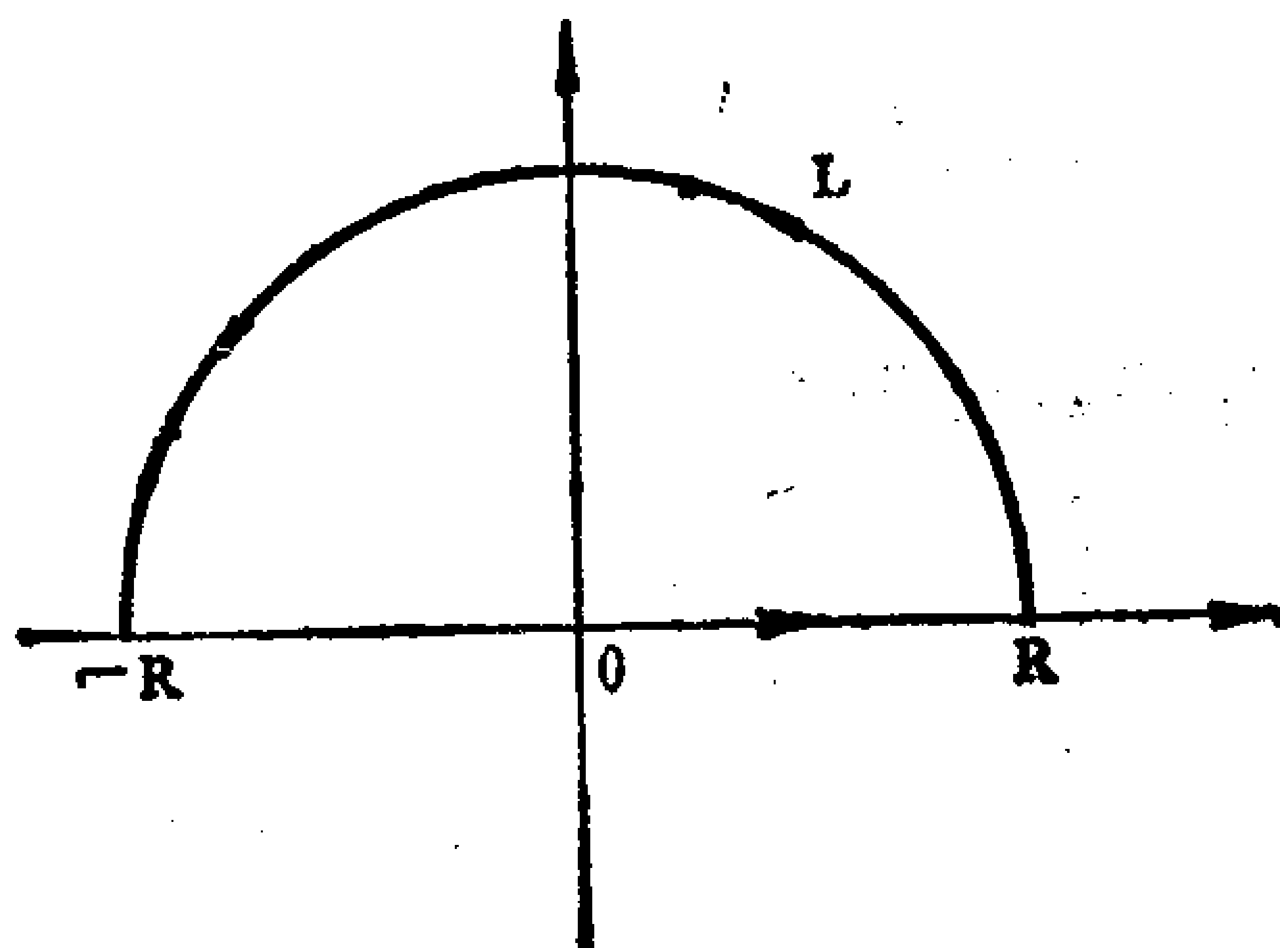


图5—1

故有 $$\int_L e^{itz}\frac{1}{1+z^2}dz$$

$$=\int_{-R}^{R} e^{itx}\frac{1}{1+x^2}dx+\int_0^{\pi} iRe^{i\theta}\cdot\frac{e^{itRe^{i\theta}}}{1+R^2e^{i2\theta}}d\theta$$

$$\triangleq I_1+I_2. \tag{2}$$

对其中I_1有

$$\lim_{R\to+\infty} I_1=\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx. \tag{3}$$

由于 $t>0$，$\theta\in[0,\pi]$，$0<e^{-tR\sin\theta}\leqslant 1$，

所以有

$$|I_2|\leqslant R\int_0^{\pi}\left|\frac{e^{itRe^{i\theta}}}{1+R^2e^{2\theta i}}\right|d\theta$$

$$=R\int_0^{\pi}\left|\frac{e^{-tR\sin\theta}\cdot e^{itR\cos\theta}}{1+R^2e^{2\theta i}}\right|d\theta$$

$$\leqslant \frac{R}{R^2-1}\int_0^{\pi} e^{-tR\sin\theta}d\theta$$

$$\leqslant \frac{R\pi}{R^2-1} \to 0 \quad (R \to +\infty). \tag{4}$$

在上半平面仅有$z=i$是(**)中被积函数$e^{itz}\dfrac{1}{1+z^2}$ 的一阶奇点。由留数定理知

$$\int_L e^{itz}\frac{1}{1+z^2}dz = 2\pi i\cdot \text{Res}\left(e^{itz}\frac{1}{1+z^2},\ i\right)$$

$$= 2\pi i \lim_{z\to i}\frac{e^{itz}}{1+z^2}(z-i)$$

$$= 2\pi i \lim_{z\to i}\frac{e^{itz}}{(z+i)}$$

$$= 2\pi i\cdot\frac{e^{-t}}{2i} = \pi e^{-t}. \tag{5}$$

综合(1)、(2)、(3)、(4)、(5)，当$t>0$时，得

$$\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx = \pi e^{-t} = \pi e^{-|t|}. \tag{6}$$

当$t<0$时，

$$\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx \quad (令x=-y)$$

$$= \int_{-\infty}^{\infty} e^{-ity}\frac{1}{1+y^2}dy$$

$$= \pi e^{-(-t)} = \pi e^{t} = \pi e^{-|t|}. \tag{7}$$

当$t=0$时，

$$\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx = \int_{-\infty}^{\infty}\frac{1}{1+x^2}dx = \pi.$$

由(6)、(7)、(8)得

$$\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx=\pi e^{-|t|}\quad(-\infty<t<\infty).$$

于是 $\varphi_\xi(t)=e^{-|t|}$.

同理可得$\varphi_\eta(t)=e^{-|t|}$,

而 $\varphi_{\xi+\eta}(t)=\varphi_{2\xi}(t)=\varphi_\xi(2t)$

$$=e^{-|2t|}=e^{-2|t|}=\varphi_\xi(t)\varphi_\eta(t).$$

但是ξ和η不独立。这是因为可取c，使$0<P(\xi<c)<1$，于是

$$P(\xi<c,\ \eta<c)=P(\eta<c)$$
$$\neq P(\xi<c)\cdot P(\eta<c).$$

进一步的讨论

我们还可以引出多元特征函数概念及其性质。利用这种性质可以构造本款的另一个反例。

1 若随机向量$(\xi_1,\ \xi_2,\ \cdots,\ \xi_n)$的分布函数为$F(x_1,x_2,\cdots,\ x_n)$。定义它的特征函数

$$\phi(t_1,t_2,\cdots,t_n)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{i(t_1x_1+\cdots+t_nx_n)}\cdot dF(x_1,\ \cdots,\ x_n).$$

2 如果$\phi(t_1,\ t_2,\ \cdots,\ t_n)$是$(\xi_1,\ \xi_2,\ \cdots,\ \xi_n)$的特征函数，则

$$\eta=a_1\xi_1+a_2\xi_2+\cdots+a_n\xi_n$$

的特征函数

$$\phi_\eta(t)=\phi(a_1t,\ a_2t,\ \cdots,\ a_nt).$$

3 考虑二维随机向量$(\xi_1,\ \xi_2)$，其分布密度可表成这样的形式

$$f(x,\ y)=f_1(x)f_2(y)+g_1(x)g_2(y)-g_1(y)g_2(x).$$

其中$f_1(x)$和$f_2(y)$是一维分布密度函数，而$g_1(x)$和$g_2(x)$ (g_1

$(x) \neq g_2(x)$)是可积分的奇函数(容易看出，这样的二维分布密度函数是存在的，例如，函数

$$f(x,y)=\frac{1}{4}e^{-|x|-|y|}\{1+xye^{-2|x|-|y|}-xye^{-|x|-2|y|}\}$$

就是这种例子)。

对任何(x, y)，不难验证

$$f(x, y)>0$$

和
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x, y)dxdy=1.$$

分量ξ_1的分布密度为

$$\begin{aligned}f_{\xi_1}(x)&=\int_{-\infty}^{\infty}f(x, y)dy\\&=\int_{-\infty}^{\infty}[f_1(x)f_2(y)+g_1(x)g_2(y)\\&\quad-g_1(y)g_2(x)]dy\\&=f_1(x).\end{aligned}$$

而分量ξ_2的分布密度为

$$f_{\xi_2}(y)=f_2(y).$$

因为 $f(x, y)\neq f_1(x)f_2(y)$,

所以ξ_1和ξ_2不独立．但是，它们的特征函数之间却有关系

$$\varphi_{\xi_1+\xi_2}(t)=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t).$$

事实上，随机向量(ξ_1, ξ_2)的二元特征函数

$$\begin{aligned}\phi(t_1, t_2)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}f(x,y)dxdy\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}\\&\quad\cdot[f_1(x)f_2(y)+g_1(x)g_2(y)\\&\quad-g_1(y)g_2(x)]dxdy\end{aligned}$$

$$
\begin{aligned}
&=\left(\int_{-\infty}^{\infty}e^{it_1x}f_{\xi_1}(x)dx\right)\left(\int_{-\infty}^{\infty}e^{it_2y}f_{\xi_2}(y)dy\right)\\
&\quad+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}g_1(x)g_2(y)dxdy\\
&\quad-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}g_1(y)g_2(x)dxdy\\
&=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t)+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}\\
&\quad\cdot g_1(x)g_2(y)dxdy\\
&\quad-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{it_1x+it_2y}g_1(y)g_2(x)dxdy.
\end{aligned}
$$

由此可求得$\xi_1+\xi_2$的特征函数

$$
\begin{aligned}
\varphi_{\xi_1+\xi_2}(t)&=\phi(t,\ t)\\
&=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t)+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{itx+ity}\\
&\quad\cdot g_1(x)g_2(y)dxdy-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{itx+ity}\\
&\quad\cdot g_1(y)g_2(x)dxdy\\
&=\varphi_{\xi_1}(t)\cdot\varphi_{\xi_2}(t).
\end{aligned}
$$

5—2 当k为奇数时，随机变量ξ的特征函数$\varphi(t)$在$t=0$处可微分k次，但$E\xi^k$不存在的例子

有关概念和命题

1 设随机变量ξ的n阶矩存在，则特征函数$\varphi(t)$可微分k次($k\leqslant n$)，且

$$\varphi^{(k)}_{(0)}=i^kE\xi^k$$

但是，当n为奇数时，以上性质的逆是不成立的，也就是说特征

函数在0点可微分k次，但k阶矩未必存在。

在举反例以前，需要证明一个引理：

2　引理　如果随机变量的密度函数$f(x)$为偶函数，则它的特征函数$\varphi(t)$取实值，且

$$-1\leqslant\varphi(t)\leqslant 1.$$

事实上，利用

$$f(x)=f(-x)$$

及

$$e^{ix}=\cos x+i\sin x,$$

可得

$$\begin{aligned}\varphi(t)&=\int_{-\infty}^{0}e^{itx}f(x)dx+\int_{0}^{+\infty}e^{itx}f(x)dx\\&=\int_{0}^{+\infty}(e^{-itx}+e^{itx})f(x)dx\\&=2\int_{0}^{+\infty}\cos tx f(x)dx.\end{aligned}$$

因而$\varphi(t)$取实值，且$-1\leqslant\varphi(t)\leqslant 1$。

反例

设随机变量ξ的密度函数

$$f(x)=\begin{cases}0, & |x|\leqslant 2,\\ \dfrac{c}{x^2\log|x|}, & |x|>2.\end{cases}$$

其中c为常数，由条件$\int_{-\infty}^{\infty}f(x)dx=1$决定。由于$f(-x)=f(x)$，由引理可知$\xi$的特征函数

$$\varphi(t)=2c\int_{2}^{\infty}\frac{\cos tx}{x^2\log x}dx.$$

由此可知，$\dfrac{1-\varphi(t)}{2c}$是t的非负，实值偶函数。利用不等式$0\leqslant 1-\cos x\leqslant\min(2,\ x^2)$，对$t<\dfrac{1}{2}$得

$$0\leqslant\frac{1-\varphi(t)}{2c}=\int_2^{\infty}\frac{1-\cos tx}{x^2\log x}dx$$

$$=\int_2^{\frac{1}{t}}\frac{1-\cos tx}{x^2\log x}dx+\int_{\frac{1}{t}}^{\infty}\frac{1-\cos tx}{x^2\log x}dx$$

$$\leqslant 2t^2\int_2^{\frac{1}{t}}\frac{dx}{\log x}+2\int_{\frac{1}{t}}^{\infty}\frac{dx}{x^2\log x}.$$

因为 $$2t^2\int_2^{\frac{1}{t}}\frac{dx}{\log x}=0(t),$$

$$2\int_{\frac{1}{t}}^{\infty}\frac{dx}{x^2\log x}=0(t),$$

所以 $$0\leqslant\frac{1-\varphi(t)}{2c}\leqslant 0(t).$$

故当$t>0$时，有

$$0\leqslant\frac{1-\varphi(t)}{t}\leqslant\frac{2c\cdot 0(t)}{t}=0(1).$$

由此知 $$\lim_{t\to 0^+}\frac{1-\varphi(t)}{t}=0.$$

注意到$\varphi(0)=1$，故得$\varphi'(0+0)=0$. 又因为

$$\lim_{t\to 0^-}\frac{\varphi(t)-1}{t}=\lim_{s\to 0^+}\frac{\varphi(-s)-1}{-s}$$

$$=\lim_{s\to 0^+}\frac{1-\varphi(s)}{s}=0,$$

故$\varphi'(0-0)=\varphi'(0+0)$，从而$\varphi'(0)=0$.

但是，因为

$$E|\xi|=\int_{-\infty}^{\infty}|x|f(x)dx=2\int_2^{+\infty}x\cdot\frac{C}{x^2\log x}dx$$

$$=2c\int_2^{+\infty}\frac{1}{x\log x}dx=2C\cdot\lim_{A\to+\infty}\int_2^{A}\frac{1}{x\log x}dx$$

$$= 2c \lim_{A\to+\infty} (\log\log A - \log\log 2) = +\infty.$$

可知一阶矩$E\xi$不存在。

进一步的讨论

关于随机变量的特征函数与矩的关系，综合起来有如下的重要结果。

定理 设F是一个分布函数，具有直到n阶的各阶有限矩，则F的特征函数φ有直到n阶的连续导数，且关系式

$$\varphi^{(k)}_{(0)} = i^k \cdot E\xi^k.$$

对$k=1, 2, \cdots, n$成立，而且φ有展式：

$$\varphi(t) = 1 + \sum_{k=1}^{n} (E\xi^k) \frac{(it)^k}{k!} + 0(t^n) \quad (t\to 0).$$

反之，假定一个分布函数F的特征函数φ有下列展式

$$\varphi(t) = 1 + \sum_{k=1}^{n} (a_k) \frac{(it)^k}{k!} + 0(t^n) \quad (t\to 0),$$

则当n为偶数时，F有直到n阶的有限矩，且

$$E\xi^k = a_k \quad (k\leqslant n).$$

当n为奇数时，F有直到$n-1$阶的矩，且

$$E\xi^k = a_k \quad (k\leqslant n-1)$$

（证明参阅〔1〕，P142）。

5—3 分布函数绝对连续，但其对应的特征函数不绝对可积的例子

有关概念和命题

定理 设特征函数$\varphi(t)$绝对可积，即

$$\int_{-\infty}^{\infty}|\varphi(t)|dt<\infty,$$

则对应的分布函数$F(x)$是连续型的，$F'(x)$处处存在，有界而且连续，又对一切$x\in R_1$，

$$F'(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\varphi(t)dt.$$

但是$\varphi(t)$的绝对可积性并非必要，即使它不成立，定理结论仍可能正确。请看下述反例：

反例

设$\varphi(t)=(1+|t|)^{-1}$，则不难验证它满足Pólya 定理（见〔1〕P168定理3.6.3）的四个条件，故$\varphi(t)$是一个绝对连续的分布函数所对应的特征函数。然而，$\varphi(t)$却是不绝对可积的，因为

$$\int_{-\infty}^{\infty}\frac{1}{1+|t|}dt=2\int_{0}^{\infty}\frac{1}{1+t}dt=\infty.$$

5—4 特征函数$\varphi(t)$在有限区间内的值不足以唯一确定此$\varphi(t)$，从而也不足以唯一决定分布函数$F(x)$的例子

有关概念和命题

1 两个分布函数$F_1(x)$和$F_2(x)$恒等的充要条件是它们各自的特征函数$\varphi_1(t)$和$\varphi_2(t)$恒等。

2 设特征函数$\varphi(t)$绝对可积，即

$$\int_{-\infty}^{\infty}|\varphi(t)|dt<\infty,$$

则对应分布函数是连续型的，且对一切$x\in R_1$，

$$F'(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\varphi(t)dt.$$

下面的例子表明，一特征函数在有限区间上的值不足以唯一决定此特征函数，因而也不足以唯一决定分布函数。

反例

(Гнеgенко)令特征函数

$$\varphi_1(t)=\begin{cases}1-|t|, & |t|\leqslant 1,\\ 0, & |t|>1.\end{cases} \tag{1}$$

由于 $\varphi_1(t)$在R_1上绝对可积，故对应密度函数为

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\varphi_1(t)dt$$

$$=\frac{1}{2\pi}\int_{-1}^{0}(1+t)e^{-itx}dt+\frac{1}{2\pi}\int_{0}^{1}(1-t)e^{-itx}dt$$

$$=\frac{1}{2\pi}\left[-\frac{1}{ix}-\frac{1}{(ix)^2}(1-e^{ix})\right]$$

$$+\frac{1}{2\pi}\left[\frac{1}{ix}+\frac{1}{(ix)^2}(e^{-ix}-1)\right]$$

$$=\frac{1}{\pi x^2}\left(1-\frac{e^{ix}+e^{-ix}}{2}\right)=\frac{1-\cos x}{\pi x^2}. \tag{2}$$

现考虑一离散型分布，它的概率分布为

$$P(\xi=0)=\frac{1}{2},\ P(\xi=(2k-1)\pi)=\frac{2}{(2k-1)^2\pi^2}$$

$$(k=0,\ \pm 1,\ \pm 2,\ \cdots). \tag{3}$$

这分布的特征函数为

$$\varphi_2(t)=\frac{1}{2}+\sum_{k=-\infty}^{+\infty}\frac{2}{(2k-1)^2\pi^2}e^{it(2k-1)\pi}$$

$$=\frac{1}{2}+\frac{2}{\pi^2}\sum_{k=-\infty}^{\infty}\frac{\cos(2k-1)t\pi+i\sin(2k-1)t\pi}{(2k-1)^2}$$

$$=\frac{1}{2}+\frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(2k-1)t\pi}{(2k-1)^2}.$$

以下来证当$|t|\leqslant 1$时，$\varphi_1(t)=\varphi_2(t)$。

实际上，在区间$|t|\leqslant 1$上把函数$g(t)=|t|$展成以2为周期的傅立叶级数

$$g(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}a_n\cos n\pi t,$$

其中系数$\frac{a_0}{2}=\int_0^1 t\,dt=\frac{1}{2}$

$$\begin{aligned}a_n&=2\int_0^1 t\cdot\cos n\pi t\,dt\\&=\left[\frac{2t\sin n\pi t}{n\pi}\right]_0^1-\frac{2}{n\pi}\int_0^1\sin n\pi t\,dt\\&=-\frac{2}{n\pi}\left[-\frac{\cos n\pi t}{n\pi}\right]_0^1=2\frac{\cos n\pi-1}{n^2\pi^2}\\&=2\cdot\frac{(-1)^n-1}{n^2\pi^2}.\end{aligned}$$

因而　　$a_{2k}=0\ (k\neq 0)$，　$a_{2k-1}=-\frac{4}{(2k-1)^2\pi^2}$，

于是　　$$g(t)=|t|=\frac{1}{2}-\frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(2k-1)\pi t}{(2k-1)^2}. \tag{5}$$

代入(1)，将所得结果与(4)比较，可见当$|t|\leqslant 1$时，$\varphi_1(t)=\varphi_2(t)$。但是$\varphi_1(t)$和$\varphi_2(t)$是两个不同分布（一个是连续型分布，另一个是离散型分布）的特征函数。

进一步的讨论

寻求特征函数被其在一个区间（$-T$，T）上的值完全确定

的条件．下面的定理（Marcinkiewicz）给出了这个问题的部分解答。

定理 设给定一个特征函数

$$\varphi(t)=\int_{-\infty}^{\infty}e^{itx}dF(x).$$

（1）如果$\int_{-\infty}^{\infty}e^{izx}dF(x)$在$z_0$点收敛，则$\int_{-\infty}^{\infty}e^{izx}dF(x)$是$\varphi(t)$在实轴与过$z_0$的水平线之间的条形区域$D$上的解析开拓。

（2）如果$\varphi(t)$可从原点邻近的一个区间$(-T, T)$出发，解析开拓到复平面的一侧或两侧，且从原点循虚轴向上及向下行走时首先遇到的自然边界点分别为A与B，则$\phi(t)$至少可以解析开拓到过A与B的两条水平线之间的条形区域，且在此区域内其解析开拓必定就是$\int e^{izx}dF(x)$。

（3）在（2）的假定下，$\varphi(t)$由其在$(-T, T)$上的值完全确定（证明参阅〔3〕，中译本P228—P229）。

5—5 特征函数列的极限函数不是特征函数的例子

有关概念

特征函数$\varphi(t)$在R_1上一致连续，且

$$|\varphi(t)|\leqslant\varphi(0)=1.$$

反例

设 $$\varphi_n(t)=\begin{cases}\dfrac{\sin nt}{nt}, & t\neq 0,\\ 1, & t=0,\end{cases}\qquad (n=1,2,\cdots)$$

则$\{\varphi_n(t)\}$是一列特征函数．事实上

$$\varphi_n(t)=\frac{1}{2n}\int_{-n}^{n}\cos tx dx=\frac{1}{2n}\int_{-n}^{n}e^{itx}dx$$
$$=\int_{-\infty}^{\infty}e^{itx}f_n(x)dx.$$

其中 $$f_n(x)=\begin{cases}\frac{1}{2n}, & x\in[-n,\ n],\\ 0, & x\bar{\in}[-n,\ n].\end{cases}$$

是分布函数

$$F_n(x)=\begin{cases}0, & x<-n,\\ \frac{x+n}{2n}, & -n\leqslant x<n,\\ 1, & x\geqslant n\end{cases}$$

的密度函数．显然，对每个 $t\in R_1$，有$\lim\limits_{n\to\infty}\varphi_n(t)=\varphi(t)$，这里

$$\varphi(t)=\begin{cases}1, & t=0,\\ 0, & t\neq 0.\end{cases}$$

因为$\varphi(t)$在$t=0$处不连续，故它不可能为特征函数。

进一步的讨论

附带指出，对上例中的分布函数列$F_n(x)$，其极限函数

$$F(x)=\lim_{n\to\infty}F_n(x)=\frac{1}{2}\qquad(\forall x_1\in R_1)$$

它不是一个分布函数．其理由是显然的。

5—6 分布函数不具有再生性的例子

有关概念和命题

1 （卷积公式）设ξ_1和ξ_2独立，则和$\xi_1+\xi_2$的分布密度

$$f_{\xi_1+\xi_2}(z)=\int_{-\infty}^{\infty}f_{\xi_1}(x)f_{\xi_2}(z-x)dx \qquad (1)$$

或

$$f_{\xi_1+\xi_2}(z)=\int_{-\infty}^{\infty}f_{\xi_2}(y)f_{\xi_1}(z-y)dy. \qquad (2)$$

其中f_{ξ_1}、f_{ξ_2}分别为ξ_1和ξ_2的分布密度函数。公式(1)或(2)称为卷积公式。

2（分布函数的再生性）当两个独立的随机变量的分布属于同一类型的分布时，如果它们的和也服从这种类型的分布，则称这种分布具有再生性。

正态分布、Poisson分布、二项分布、χ^2—分布等都具有再生性。但是，也确实存在某种分布，它不具有再生性。

反例

设随机变量ξ_1和ξ_2独立，并且都在区间$[-a, a]$上服从均匀分布，易知ξ_1和ξ_2的分布密度函数分别是

$$f_{\xi_1}(x)=\begin{cases}\dfrac{1}{2a}, & |x|\leqslant a,\\ 0, & |x|>a\end{cases}$$

和

$$f_{\xi_2}(y)=\begin{cases}\dfrac{1}{2a}, & |y|\leqslant a,\\ 0, & |y|>a.\end{cases}$$

由卷积公式得

$$f_{\xi_1+\xi_2}(z)=\int_{-a}^{a}\frac{1}{2a}f_{\xi_2}(z-x)dx.$$

作变量替换$z-x=t$，得到

$$f_{\xi_1+\xi_2}(z)=\frac{1}{2a}\int_{z-a}^{z+a}f_{\xi_2}(t)dt.$$

现分四种情况来讨论：

(1) 当$z<-2a$时，积分限$z-a$及$z+a$都小于$-a$，被积函数在积分区间上为0，因而

$$f_{\xi_1+\xi_2}(z)=0.$$

(2) 当$z>2a$时，积分限$z-a$及$z+a$都大于a，类似地有

$$f_{\xi_1+\xi_2}(z)=0.$$

(3) 当$-2a\leqslant z\leqslant 0$时，我们有

$$-3a\leqslant z-a<-a,\quad -a\leqslant z+a\leqslant a$$

把积分区间分成子区间$[z-a,\ -a]$及$[-a,\ z+a]$，被积函数在第一个区间上为0，而在第二个区间上等于$\frac{1}{2a}$，所以得到

$$f_{\xi_1+\xi_2}(z)=\frac{1}{4a^2}\int_{-a}^{z+a}dt=\frac{2a+z}{4a^2}.$$

(4) 当$0\leqslant z\leqslant 2a$时，我们有

$$-a\leqslant z-a\leqslant a,\quad a\leqslant z+a\leqslant 3a.$$

把积分区间分成子区间$[z-a,\ a]$及$[a,\ z+a]$，则被积函数在第一个区间上等于$\frac{1}{2a}$，而在第二个区间上为0，所以得到

$$f_{\xi_1+\xi_2}(z)=\frac{1}{4a^2}\int_{z-a}^{a}dt=\frac{2a-z}{4a^2}.$$

综上所述，我们得到随机变量$\xi_1+\xi_2$的分布密度为

$$f_{\xi_1+\xi_2}(z)=\begin{cases}\dfrac{4a+z}{4a^2}, & 当-2a\leqslant z\leqslant 0,\\ \dfrac{2a-z}{4a^2}, & 当0\leqslant z\leqslant 2a,\\ 0, & 当|z|>2a.\end{cases}$$

所得分布叫做辛普生分布。

进一步的讨论

还有人研究了分布函数再生性这类命题的逆命题——分布函数的分解问题。即若两个独立随机变量之和服从某一分布，问是否能断定这两个随机变量也都服从这个分布。已经证明，

对于正态分布、Poisson分布和二项分布，逆命题是成立的。它们的证明分别由H. 克拉梅、Д. A拉依科夫和D. N. 森伯哈等给出。

5—7 分布函数$F(x)$不是无穷可分分布的例子

有关概念和命题

1 一个分布函数$F(x)$称为无穷可分的，如果对每一个自然数n，都存在一个分布函数F_n，使得

$$F = \overbrace{F_n * F_n * \cdots * F_n}^{n\text{次}}.$$

这里$F_1 * F_2$表示F_1和F_2的卷积。

这个定义等价于：一个特征函数φ称为是无穷可分的，如果对每一个自然数n，存在一个特征函数φ_n，使得

$$\varphi = [\varphi_n]^n.$$

正态分布、Poisson分布、Gamma分布都是无穷可分的。这一点留给读者自行证明。

2 若$\varphi(t)$是无穷可分的特征函数，则$\varphi(t)$不会等于0（证明参阅〔18〕，P279—P280）。

反例

设随机变量ξ服从$[-1, 1]$上的均匀分布，则其特征函数$\varphi(t)=\frac{\sin t}{t}$。因为对某些$t$、$\varphi(t)$可等于0，所以它不可能是无穷可分的。

5—8 无处为0的特征函数不是无穷可分的例子

有关概念 见例5—7。

我们已经知道，若$\varphi(t)$是无穷可分的特征函数，则$\varphi(t)$无处为0的（即$\varphi(t)$无实0点），但其逆不真，我们有如下的反例：

反例

设 $P(\xi=0)=1-P$，$p(\xi=1)=p$。

记$q=1-p$，则ξ的特征函数为

$$\varphi(t)=(q+pe^{it}).$$

因为$\varphi(t)\neq 0$，对一切$t\in R_1$，即$\varphi(t)$是无处为0的。但由于$\varphi(t)$是不可分解的，由此可知$\varphi(t)$不可能是无穷可分的。

5—9 无穷可分的特征函数可以分解为不是无穷可分的特征函数的乘积的例子

有关概念和命题

1 关于无穷可分的特征函数的概念见例5—7。

2 定理（Lèvy-Khintchine表达式）一个定义在R_1上的复值函数φ是无穷可分特征函数的充要条件是$\ln\varphi(t)$能有表达式

$$\ln\varphi(t)=i\alpha t+\int_{-\infty}^{\infty}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}dG(x).$$

其中$\alpha\in R_1$且G是R_1上的有界、非降、右连续函数，满足$G(-\infty)=0$和$G(+\infty)<\infty$。在此处被积函数在$x=0$处的值由连续

性定义为

$$\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}\bigg|_{x=0}=-\frac{t^2}{2}.$$

此外，α和G由φ唯一确定，其中G被称为φ的Lévy谱（证明参阅〔1〕P.239—244）。

3　定理　有限多个无穷可分特征函数的乘积仍是无穷可分的（见〔1〕，P235）。但其逆不真，即无穷可分的特征函数不一定是无穷可分的特征函数的乘积。请看下面的反例

反例

设$0<a\leqslant b<1$，考虑复值函数

$$\varphi(t)=\frac{1-b}{1-a}\ \frac{1+ae^{-it}}{1-be^{it}}.$$

因为$\varphi(t)$在$t=0$点连续，且

$$\varphi(t)=\frac{1-b}{1+a}\left[ae^{-it}+(1+ab)\sum_{k=0}^{\infty}b^ke^{itk}\right],$$

所以$\varphi(t)$是随机变量ξ的特征函数，其中ξ具有分布

$$P(\xi=-1)=\frac{1-b}{1+a}a,\ P(\xi=k)=\frac{1-b}{1+a}(1+ab)b^k$$

$$(k=0,\ 1,\ 2,\ \cdots).$$

由$\varphi(t)$的定义可知

$$\ln\varphi(t)=\sum_{k=1}^{\infty}\left[(-1)^{k-1}\frac{a^k}{k}(e^{-ikt}-1)+\frac{b^k}{k}(e^{itk}-1)\right],$$

于是取　$\alpha=\sum_{k=1}^{\infty}\frac{b^k+(-1)^ka^k}{k^2+1}$,

而G是一个有界变差函数，它在

$$x=k\text{处有跳跃}\frac{kb^k}{k^2+1}$$

和 $$x=-k\text{处有跳跃}(-1)^{k-1}\frac{ka^k}{k^2+1}\quad(k=1,2,3,\cdots),$$

则 $$\ln\varphi(t)=it\cdot\sum_{k=1}^{\infty}\frac{b^k+(-1)^k a^k}{k^2+1}+\sum_{k=1}^{\infty}\left\{\left[\left(e^{itk}-1-\frac{itk}{1+k^2}\right)\frac{1+k^2}{k^2}\cdot\frac{kb^k}{k^2+1}\right]+\left[\left(e^{-itk}-1-\frac{-itk}{1+k^2}\right)\frac{1+k^2}{k^2}\quad\frac{(-1)^{k-1}\cdot ka^k}{k^2+1}\right]\right\},$$

即$\ln\varphi(t)$能表成Lévy-Khintchine表达式的形式。然而 $G(x)$显然不是单调的，所以$\varphi(t)$不是无穷可分的。而且注意到$\bar{\varphi}$ 也不是无穷可分的，但是$|\varphi|$是无穷可分的。由于

$$\bar{\varphi}(t)=\frac{1-b}{1+a}\quad\frac{1+ae^{it}}{1-be^{-it}},$$

于是 $$\ln\overline{\varphi(t)}=\sum_{k=1}^{\infty}\left[\frac{b^k}{k}[e^{-ikt}-1]+(-1)^{k+1}\frac{a^k}{k}(e^{ikt}-1)\right].$$

令 $$\theta(t)=|\varphi(t)|^2=\varphi(t)\ \bar{\varphi}(t),$$

于是 $$\ln\theta(t)=\sum_{k=1}^{\infty}\frac{1}{k}[b^k+(-1)^{k-1}a^k](e^{-ikt}-1)+\sum_{k=1}^{\infty}\frac{1}{k}[b^k+(-1)^{k-1}a^k](e^{ikt}-1).$$

取$a=0$，G是非降函数，且在

$$\xi = \pm k \text{处有跃度} \frac{k}{k^2+1}[b^k + (-1)^{k-1}a^k]$$

$$(k=1, 2, 3, \cdots),$$

则知θ有唯一的Lévy-Khintchine表达式，所以θ是无穷可分的。

这表明无穷可分的特征函数$\theta(t)$可表为二个不是无穷可分的特征函数$\varphi(t)$与$\bar{\varphi}(t)$的乘积。

5—10 随机变量的矩母函数不存在的例子

有关概念

若ξ是定义在$(\Omega, \mathscr{F}, P)$上的随机变量，函数

$$M(s) = Ee^{s\xi} \qquad (*)$$

称为ξ的矩母函数，如果（*）右方的期望在原点的某个邻域内存在。

反例

例1 设随机变量ξ有概率密度

$$P(\xi = k) = \frac{6}{\pi^2} \cdot \frac{1}{k^2} \qquad (k=1, 2, \cdots),$$

因为
$$\sum_{k=1}^{\infty} e^{sk} \cdot P(\xi = k) = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{e^{sk}}{k^2}$$

对每一个$s > 0$都是发散的，所以ξ的矩母函数不存在。

例2 哥西分布的矩母函数也不存在。事实上，哥西分布的密度函数

$$f(x) = \frac{1}{\pi^2} \frac{1}{1+x^2} \qquad (x \in R_1),$$

故
$$E|e^{s\xi}| = \int_{-\infty}^{\infty} e^{sx} \cdot \frac{1}{\pi^2} \frac{1}{1+x^2} dx = \frac{1}{\pi^2} \int_{-\infty}^{\infty} \frac{e^{sx}}{1+x^2} dx$$

对任何$s>0$都发散。

5—11 随机变量ξ的各阶矩都存在，但矩母函数不存在的例子

有关概念 见例5—10

反例

设随机变量ξ的密度函数

$$f(x)=Ce^{-|x|^{\alpha}} \qquad 0<\alpha<1,\ x\in R_1.$$

其中常数C由下式决定

$$C\int_{-\infty}^{\infty}e^{-|x|^{\alpha}}dx=1.$$

设$s>0$，则

$$\int_0^{\infty}e^{sx}e^{-|x|^{\alpha}}dx=\int_0^{\infty}e^{x(s-x^{\alpha-1})}dx.$$

因为$\alpha-1<0$，故$\int_0^{\infty}e^{sx}e^{-x^{\alpha}}dx$对任何$s>0$都发散。

由于 $$\int_{-\infty}^{\infty}e^{sx}e^{-|x|^{\alpha}}dx=\int_{-\infty}^{0}e^{sx}e^{-(-x)^{\alpha}}dx+\int_0^{+\infty}e^{sx}e^{-x^{\alpha}}dx$$
$$\geqslant\int_0^{+\infty}e^{sx}e^{-x^{\alpha}}dx$$

对任何$s>0$都成立，所以ξ的矩母函数不存在。

但是 $$E|\xi|^n=c\int_{-\infty}^{\infty}|x|^n e^{-|x|^{\alpha}}dx$$

注1 ξ之矩母函数$Ee^{s\xi}$存在之充要条件为
(i) ξ之k阶矩$\alpha_k=E\xi^k$存在（$K\geqslant1$）
(ii) 存在正数v使
$|\alpha_k|\leqslant K!v^k$（$K\geqslant1$）
（见〔1〕P253之定理4.2.2和P254之注4.2.2）

$$= 2c\int_0^{+\infty} x^n e^{-x^\alpha} dx$$

$$= 2c\int_0^{+\infty} y^{n/\alpha} e^{-y} \cdot \frac{1}{2} y^{\frac{1}{\alpha}-1} dy$$

$$= \frac{2c}{\alpha}\int_0^{+\infty} y^{\frac{n+1}{\alpha}-1} e^{-y} dy$$

对每个n都存在，即$E\xi^n$都存在

(*) 5—12 并非所有的特征函数都是解析的例子

定义 设$\varphi(t)$是实变量的实值特征函数，$\theta(z)$是复变量的复值函数，若$\theta(z)$在$|z|<\rho(\rho>0)$上全纯（或正则）并且存在$\delta>0$，使在$|t|<\delta$上有

$$\varphi(t) = \theta(z),$$

则称$\varphi(t)$是一个解析的特征函数。

我们知道，二项、Poisson、正态和Gamma等分布所对应的特征函数都是解析的特征函数，但并非所有的特征函数都是解析的。请看

例1 设

$$\xi \sim f(x) = \frac{1}{\pi}\frac{\alpha}{\alpha^2 + x^2} \qquad (\alpha>0),$$

则ξ之特征函数

$$\varphi(t) = e^{-|t|^\alpha} \qquad (0<\alpha<2).$$

但它不是解析的。事实上，由于$\varphi(t)$在$t=0$处的二阶导数不存在，故其二阶矩不存在，于是由文献〔1〕P253之知理4.2.2知

$$\varphi(t) = e^{-|t|^\alpha} \qquad (0<\alpha<2)$$

不是解析的。

(*) 5—13　消去法对一般特征函数之分解不一定成立的例子

例1　设特征函数

$$\varphi_1(t)=\begin{cases}1-|t|, & \text{当}|t|\leqslant 1,\\ 0, & \text{当}|t|>1.\end{cases}$$

$$\varphi_2(t)=1-|t-2k|,\ \text{当}2k-1\leqslant t\leqslant 2k+1,\ k=0,\ \pm 1,\ \pm 2,\ \cdots$$

我们有　$\varphi_2(t+2)=\varphi_2(t)$，

即$\varphi_2(t)$是周期为2的特征函数. 显然

$$\varphi_1(t)=\varphi_2(t)\quad \text{当}|t|\leqslant 1\text{时},$$

且　$\varphi(t)\triangleq\varphi_1(t)\varphi_2(t)=\varphi_1(t)\varphi_1(t)$，对$t\in R_1$成立.

但　$\varphi_1(t)\neq\varphi_2(t)$，对$t\in R_1$.

即消去法对一般特征函数之分解并不一定成立.

注　然而若φ、φ_1、φ_2皆是无穷可分的特征函数，且

$$\varphi=\varphi_1\varphi_2=\varphi_1\varphi_3$$

时（φ_3为一般的特征函数，不必是无穷可分的），则由无穷可分的特征函数$\varphi_1(t)$无0点，可得$\varphi_2=\varphi_3$，即上式之φ_1可在等式两边消去，即此时消去法成立.

(*) 5—14　无穷可分的特征函数存在不可分解因子的例子

有关定理

首先我们有如下著名的辛钦（Khintchine）定理：无不可

分解因子的特征函数是无穷可分的(见[1], P264定理4.3.2)。

但其逆不真，即无穷可分的特征函数不一定无不可分解的因子，我们有如下的

反例

设$0<\alpha<1$，则

$$\varphi(t)=\frac{1-\alpha}{1-\alpha e^{it}}=e^{\ln\frac{1-\alpha}{1-\alpha e^{it}}}$$

$$=exp\{\ln(1-\alpha)-\ln(1-\alpha e^{it})\}$$

$$=exp\left\{\sum_{j=1}^{\infty}\frac{\alpha^j}{j}(e^{itj}-1)\right\}$$

$$=\prod_{j=1}^{\infty}exp\left\{\frac{\alpha^j}{j}(e^{itj}-1)\right\}.$$

由于$\varphi(t)$是具有poisson型特征函数之无穷乘积，所以由文献[1]P237命题4.1.4知$\varphi(t)$是无穷可分的特征函数。

$$由于\varphi(t)=(1-\alpha)\sum_{j=0}^{\infty}(\alpha e^{it})^j$$

∴ $\varphi(t)$是几何分布

$$p(\xi=j)=(1-\alpha)\alpha^j,\ j=0,\ 1,\ 2,\ \cdots;\ 0<\alpha<1$$

的特征函数。

另一方面，对于$|x|<1$

$$(1-x)^{-1}=\sum_{j=0}^{\infty}x^j=\prod_{j=0}^{\infty}(1+x^{2^j}).$$

∴ $\varphi(t)$可改写为

$$\varphi(t)=\prod_{j=0}^{\infty}\left(\frac{1+\alpha^{2^j}e^{it2^j}}{1+\alpha^{2^j}}\right).$$

注意到 $\varphi_j(t)=\dfrac{1+a^{2^j}e^{it2^j}}{1+a^{2^j}}$

$$=\frac{1}{1+a^{2^j}}+e^{it2^j}\frac{a^{2^j}}{1+a^{2^j}},$$

∴ $\varphi_j(t)$是ξ_j之特征函数，其中ξ_j满足

$$p(\xi_j=0)=\frac{1}{1+a^{2^j}},\quad p(\xi_j=2^j)=\frac{a^{2^j}}{1+a^{2^j}},$$

且$\varphi_j(t)$对于$j=0$，1，2，…是不可分解的。即无穷可分的特征函数$\varphi(t)$它是可数个不可分解的特征函数$\varphi_j(t)$的无穷乘积。

第六章　与收敛性有关的反例

6—1　分布函数列$\{F_n(x)\}$弱收敛于$F(x)$，但$F(x)$不是分布函数的例子

有关概念

对于分布函数列$\{F_n(x)\}$，如果存在一个函数$F(x)$，使

$$\lim_{n\to\infty}F_n(x)=F(x)$$

在$F(x)$的每一个连续点上都成立，则称$F_n(x)$弱收敛于$F(x)$，并记为$F_n(x)\xrightarrow{w}F(x)$。

尽管$\{F_n(x)\}$是分布函数列，且$F_n(x)\xrightarrow{w}F(x)$，但极限函数却未必是分布函数。由此"$\{F_n(x)\}$弱收敛于分布函数$F(x)$"和"$\{F_n(x)\}$弱收敛于$F(x)$"两句话有很大的差别。前者不仅要后者成立，并且还要求$F(x)$是分布函数。

反例

例1　考虑具有退化分布的随机变量序列$\{\xi_n\}$，它的分布列为

$$p(\xi_n=n)=1\qquad(n=1,\ 2,\ 3,\ \cdots).$$

相应的分布函数

$$F_n(x)=\begin{cases}0, & x\leqslant n,\\ 1, & x>n.\end{cases}$$

令 $F(x)=0 \quad (x\in R_1)$。

显然，对任何$x\in R_1$，有

$$\lim_{n\to\infty}F_n(x)=0=F(x).$$

因而$F_n(x)\xrightarrow{w}F(x)$，但$F(x)$不是分布函数。

例2 见例5-5中的“进一步的讨论”。

6—2 分布函数列$F_n(x)\xrightarrow{w}F(x)$，但$F(x)$不唯一的例子

有关概念 见例6—1

反例

考虑正态随机变量序列$\{\xi_n\}$、$\xi_n\sim N\left(0,\frac{1}{n}\right)$它的分布函数

$$F_n(x)=\frac{n}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{n^2y^2}{2}}dy.$$

令$ny=t$，$dy=\frac{1}{n}dt$，则

$$F_n(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{nx}e^{-\frac{t^2}{2}}dt.$$

显然，当$x>0$时，$\lim\limits_{n\to\infty}F_n(x)=1$，

当$x<0$时，$\lim\limits_{n\to\infty}F_n(x)=0$，

所以 $F_n(x)\xrightarrow{w}F(x)$。

这里

$$F(x)=\begin{cases}0, & x\leqslant 0,\\ 1, & x>0.\end{cases}$$

因为$x=0$是$F(x)$的不连续点，若在$x=0$处极限函数取另外的值，例如取

$$G(x)=\begin{cases}0, & x<0,\\ \frac{1}{2}, & x=0,\\ 1, & x>0.\end{cases}$$

则仍可有$F_n(x)\xrightarrow{w}G(x)$。这个例子表明$\{F_n(x)\}$可弱收敛于两个不同的函数。

进一步的讨论

如果要求极限函数左连续，那么极限函数必唯一。即设$\{F_n(x)\}$为任一列分布函数，而且

$$F_n(x)\xrightarrow{w}G_1(x) \qquad F_n(x)\xrightarrow{w}G_2(x).$$

又$G_1(x)$和$G_2(x)$都左连续，则$G_1(x)=G_2(x)$，$(\forall x\in R_1)$。

请读者自行证明这一结果。

6—3 即使$\lim\limits_{n\to\infty}\xi_n(\omega)=\xi(\omega)$对$\forall\omega\in\Omega$都成立，也不能保证$\lim\limits_{n\to\infty}F_n(x)=F(x)$对$\forall x\in R_1$成立的例子

有关概念

在$F_n(x)\xrightarrow{w}F(x)$的定义中没有要求在极限函数$F(x)$的不连续点处的收敛性。因为如果希望$\lim\limits_{n\to\infty}F_n(x)=F(x)$对$\forall x\in R_1$都成立，那么要求就会过于苛刻，结果连处处收敛的随机变量序列也不能满足这一条件。

反例

设$C_n(n=1, 2, 3, \cdots)$为一严格上升收敛于C的常数序列，且$C<\infty$。

在概率空间$(\Omega, \mathscr{F}, p)$上定义如下的随机变量，令

$$\xi_n(\omega)=C_n \qquad (\forall\omega\in\Omega),$$

$$\xi(\omega)=C \qquad (\forall\omega\in\Omega).$$

且
$$P(\xi_n(\omega)=C_n)=1,$$

$$P(\xi(\omega)=C)=1,$$

显然，对每一个$\omega\in\Omega$，有

$$\lim_{n\to\infty}\xi_n(\omega)=\xi(\omega).$$

然而$\xi_n(\omega)$的分布函数在R_1上不是点点收敛的.

事实上，这时由于$\xi_n(n=1, 2, \cdots)$及ξ的分布函数$F_2(x)$ $(n=1, 2, \cdots)$和$F(x)$分别如下式所示：

$$F_n(x)=\begin{cases}0, & x\leqslant C_n\\ 1, & x>C_n\end{cases} \qquad (n=1, 2, \cdots).$$

$$F(x)=\begin{cases}0, & x\leqslant C,\\ 1, & x>C.\end{cases}$$

当$x\neq C$时，有$\lim\limits_{n\to\infty}F_n(x)=F(x)$。这是因为，若$x<C$，则$F(x)=0$。注意到$\{C_n\}$严格单调上升收敛于$C$，故对任何$n\geqslant1, C_n<C$，$\lim\limits_{n\to\infty}C_n=C$，因此，∃自然数$N$，当$n>N$时，$C_n\geqslant x$，于是$F_n(x)=0$。所以$\lim\limits_{n\to\infty}F_n(x)=0=F(x)$。类似地可证当$x>C$时，$\lim\limits_{n\to\infty}F_n(x)=F(x)$。但在$F(x)$的不连续点C上，$F(c)=0$，$F_n(c)=1$，故$\lim\limits_{n\to\infty}F_n(c)\neq F(c)$。

6—4 $F_n(x) \xrightarrow{w} F(x)$，但$E\xi_n^k \to E\xi^k$不成立的例子

有关概念　见例6—1。

分布函数的弱收敛推不出矩的收敛性。

反例．

设F_n为分布函数列，定义为

$$F_n(x)=\begin{cases}0, & x\leqslant 0,\\ 1-\dfrac{1}{n}, & 0<x\leqslant n,\\ 1, & x>n.\end{cases}$$

显然$F_n(x) \xrightarrow{w} F(x)$，这儿的$F(x)$是下式给定的分布函数

$$F(x)=\begin{cases}0, & x\leqslant 0,\\ 1, & x>0.\end{cases}$$

注意到$F_n(x)$是带有概率分布

$$P(\xi_n=0)=1-\frac{1}{n},\quad P(\xi_n=n)=\frac{1}{n}$$

的随机变量ξ_n的分布函数，而$F(x)$是在$x=0$处退化分布的随机变量ξ的分布函数，所以我们有

$$E\xi_n^k=n^k\cdot\frac{1}{n}=n^{k-1}\qquad (k=1,\ 2,\ \cdots),$$

$$E\xi^k=0.$$

于是对任何$k\geqslant 1$，

$$E\xi_n^k \not\longrightarrow E\xi^k.$$

进一步的讨论

在本款反例中，不仅有

$$E\xi_n^k \not\to E\xi^k,$$

而且对任何$k \geqslant 2$

$$E(\xi_n - E\xi_n)^k \not\to E(\xi - E\xi)^k.$$

这是因为$E\xi = 0$，故对任何$k \geqslant 1$.

$$E(\xi - E\xi)^k = E\xi^k = 0.$$

但 $\quad E\xi_n = 1,$

故 $\quad E(\xi_n - E\xi_n)^k = E(\xi_n - 1)^k$

$$= (-1)^k\left(1 - \frac{1}{n}\right) + (n-1)^k \cdot \frac{1}{n}.$$

所以当$k \geqslant 2$时

$$E(\xi_n - E\xi_n)^k \to \infty,$$

故 $\quad E(\xi_n - E\xi_n)^k \not\to E(\xi - E\xi)^k.$

6—5 $F_n(x) \xrightarrow{w} F(x)$，但$F_n(x) \xrightarrow{c} F(x)$不成立的例子

有关概念

若分布函数列$\{F_n\}$满足：

$$F_n(x) \xrightarrow{w} F(x),$$

且 $\quad F_n(\pm\infty) \to F(\pm\infty),$

则称$\{F_n(x)\}$完全收敛于$F(x)$，记为$F_n(x) \xrightarrow{c} F(x)$.

一般说来，$F_n(x) \xrightarrow{w} F(x)$不一定有$F_n(x) \xrightarrow{c} F(x)$.

请看下面的

反例

例1 设$F(x)=\frac{1}{2}$ $(x\in R_1)$。

对每一个$n\geqslant 1$，定义

$$F_n(x)=\begin{cases}0, & x\leqslant -n,\\ \frac{1}{2}, & -n<x\leqslant n,\\ 0 & x>n.\end{cases}$$

显然$F(x)$在整个R_1上连续，且对任何$x\in R_1$，有$F_n(x)\to F(x)$，所以$F_n(x)\xrightarrow{w}F(x)$。但当$n\to\infty$时

$$F_n(-\infty)\equiv 0 \not\to \frac{1}{2}=F(-\infty),$$

$$F_n(+\infty)\equiv 1 \not\to \frac{1}{2}=F(+\infty),$$

故 $F_n(x)\overset{c}{\not\to}F(x)$。

例2 设$G(x)$是一分布函数，则$G(-\infty)=0$，$G(+\infty)=1$。对每一个$n\geqslant 1$，定义

$$F_n(x)=G(x+n) \qquad (x\in R_1),$$

则$F_n(x)$单调不减，左连续，$F_n(-\infty)=0$，$F_n(+\infty)=1$显然对$\forall x\in R_1$

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}G(x+n)=G(+\infty)=1.$$

于是 $F_n(x)\xrightarrow{w}F(x)\equiv 1$。

但是，因为$F_n(-\infty)=0$，而$F(-\infty)=1$，所以，$F_n(x)\xrightarrow{c}F(x)$不成立。

6—6 设分布函数列$\{F_n(x)\}$对应的特征函数 $\{\varphi_n(t)\}$，若$\{\varphi_n(t)\}$收敛于某函数$\varphi(t)$， $\varphi(t)$在$t=0$不连续，则推不出$\{F_n(x)\}$弱收敛于某分布函数的例子

有关概念和命题

（逆极限定理）设特征函数列$\{\varphi_n(t)\}$收敛于某一函数$\varphi($ 且$\varphi(t)$在$t=0$连续，则相应的分布函数列$\{F_n(x)\}$弱收敛于一分布函数$F(x)$，而且$\varphi(t)$是$F(x)$的特征函数。

需要注意的是，在此定理中，$\varphi(t)$在$t=0$处的连续性是可少的条件，

反例

设

$$\varphi_n(t)=\begin{cases}\dfrac{\sin nt}{nt}, & t\neq 0\\ 1, & t=0\end{cases}\qquad (n=1,2,3\cdots),$$

则$\{\varphi_n(t)\}$是一列特征函数。事实上

$$\frac{\sin nt}{nt}=\frac{1}{2n}\int_{-n}^{n}e^{itx}dx=\int_{-\infty}^{\infty}e^{itx}f_n(x)dx.$$

其中
$$f_n(x)=\begin{cases}\dfrac{1}{2n}, & x\in[-n,n],\\ 0, & x\notin[-n,n].\end{cases}$$

是分布函数

$$F_n(x)=\begin{cases}0, & x\leqslant -n,\\ \dfrac{x+n}{2n}, & -n<x\leqslant n,\\ 1, & x>n.\end{cases}$$

的密度函数。显然，对每个t有

$$\lim_{n\to\infty}\varphi_n(t)=\varphi(t).$$

$$\varphi(t)=\begin{cases}1, & t=0,\\ 0, & t\neq 0.\end{cases}$$

t)在$t=0$处不连续，也不是特征函数。

方面，对$\forall x\in R_1$，

$$\lim_{n\to\infty}F_n(x)=F(x)\equiv\frac{1}{2},$$

$$F_n(x)\xrightarrow{w}F(x).$$

是$F(x)$不是一个分布函数。

进一步的讨论

注 由特征函数的连续性定理知，若$\varphi_n(t)\longrightarrow\varphi(t)$，$(t)$不是连续函数，则$\{F_n(x)\}$必不弱收敛于某一分布函数。

6—7 由$\xi\xrightarrow{c}\zeta$推不出相应的分布密度函数或概率分布的收敛性的例子

有关概念

1 设$\{x_i\}$为离散型随机变量的所有可能值，而$p(x_i)$是ξ取x_i的概率，即

$$P(\xi=x_i)=p(x_i)\qquad (i=1,2,3\cdots\cdots).$$

$\{p(x_i);\ i=1,2,3\cdots\}$称为随机变量$\xi$的概率分布。

2 （依分布收敛）设随机变量$\xi_n(\omega)$，$\xi(\omega)$的分布函数分别是$F_n(x)$、$F(x)$。如果$F_n(x)\xrightarrow{w}F(x)$，就称$\{\xi_n(w)\}$依分布收敛到$\xi(w)$，并记为$\xi_n(w)\xrightarrow{L}\xi(w)$。但由$\xi_n\xrightarrow{L}\xi$得不到它们

相应的密度函数或概率分布的收敛性。

反例

设$\{\xi_n\}$为离散型随机变量序列，具有概率分布

$$P_n(x) \hat{=} p(\xi_n = x) = \begin{cases} 1, & \text{当} x = 2 + \frac{1}{n}, \\ 0, & \text{其他} \end{cases}$$

注意到在点$x = 2$处，对任何$n \geq 1$，$p_n(2) = 0$，

所以 $\lim\limits_{n \to \infty} p_n(2) = 0$。

当$x \neq 2$时，若$x < 2$，则

$$\lim_{n \to \infty} p_n(x) = 0.$$

若$x > 2$，E自然数N，当$n > N$时，$2 + \frac{1}{n} < x$，故$p_n(x) = 0$，于是

$$\lim_{n \to \infty} p_n(x) = 0.$$

总之，对任何$x \in R_1$，有

$$\lim_{n \to \infty} p_n(x) \equiv 0. \qquad (*)$$

然而，ξ_n相应的分布函数为

$$F_n(x) \hat{=} p(\xi_n < x) = \begin{cases} 0, & x \leq 2 + \frac{1}{n}, \\ 1, & x > 2 + \frac{1}{n}. \end{cases}$$

显然$F_n(x) \xrightarrow{w} F(x)$，这里

$$F(x) = \begin{cases} 0, & x \leq 2, \\ 1, & x > 2. \end{cases}$$

可见$F(x)$是在$x = 2$处退化的随机变量ξ的分布函数，所以它相应的概率分布为

$$p(x)=p(\xi=x)=\begin{cases}1, & x=2,\\ 0, & 其他.\end{cases}$$

与(*)比较，知$\lim\limits_{n\to\infty}p_n(x)\neq p(x)$。此例表明由$\xi_n\xrightarrow{L}\xi$一般推不出相应概率分布的收敛性。

进一步的讨论

关于随机变量序列的依分布收敛与相应的概率分布（或密度函数）的收敛性之间的关系有如下一些结果：

定理1　设ξ_n是取整数值的随机变量序列，

令　$p_n(k)\triangleq P(\xi_n=k)$　　$(k=0,1,2\cdots)$

是$\xi_n(n=1,2,3,\cdots)$的概率分布，而

$p(k)\triangleq p(\xi=k)$

是ξ的概率分布，则

$$p_n(x)\longrightarrow p(x)\quad(x=0,1,2\cdots)\Longleftrightarrow\xi_n\xrightarrow{L}\xi$$

（证明参阅[2]，p242）。

定理2　设$\xi_n(n=1,2,3,\cdots)$和ξ是连续型随机变量，使得当$n\to\infty$时，$f_n(x)\to f(x)$几乎处处对所有$x\in R_1$成立，其中f_n和f分别为ξ_n和ξ的密度函数。则有$\xi_n\xrightarrow{L}\xi$（参阅H. Scheffé, A useful convergence theorem for probability distributions. Ann, Math, Stat, 18(1947), 434—438)。

6—8　$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$，但$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$不成立的例子

有关概念和命题

1　（依分布收敛）见6—7。

2 （依概率收敛）如果

$$\lim_{n\to\infty} P\{\omega \mid |\xi_n(\omega)-\xi(\omega)|\geq\varepsilon\}=0$$

对$\forall\varepsilon>0$成立，则称随机变量序列$\{\xi_n(\omega)\}$依概率收敛于$\xi(\omega)$，并记为$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$。

3 （性质）若$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$，则$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$，但是，当$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$时，不一定有$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$。

反例

设样本空间$\Omega=\{\omega_1,\omega_2\}$，$P(\omega_1)=P(\omega_2)=\frac{1}{2}$，定义随机变量$\xi(\omega)$如下：

$$\xi(\omega_1)=-1,\quad \xi(\omega_2)=1,$$

则$\xi(\omega)$的分布列为

$$\begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{*}$$

若对一切$n\geq1$，令$\xi_n(\omega)=-\xi(\omega)$，显然$\xi_n(\omega)$的分布列也是（*），所以$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$。但对$\varepsilon_0=1$

$$P(|\xi_n(\omega)-\xi(\omega)|>\varepsilon_0)=P(\xi_n(\omega)=1,\ \xi(\omega)=-1)$$
$$+p(\xi_n(\omega)=-1,\xi(\omega)=1)=P(\omega_1)+P(\omega_2)=1.$$

因此 $\xi_n(\omega)\overset{P}{\not\longrightarrow}\xi(\omega)$。

进一步的讨论

虽然由$\xi_n(\omega)\xrightarrow{L}\xi(\omega)$推不出$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$，但在特殊场合下却有下面的结果

定理 设C是常数，则 $\xi_n(\omega)\xrightarrow{P}C\Longleftrightarrow\xi_n(\omega)\xrightarrow{L}C$

（证明参阅[6]，p263）。

6—9　$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$但$E\xi_n^k\longrightarrow E\xi^k$（$\forall k\geqslant 1$）不成立的例子

有关概念　　见例6—8

反例

设随机变量$\xi_n(\omega)$的分布列为

$$P(\xi_n=0)=1-\frac{1}{n},\ P(\xi_n=n)=\frac{1}{n}\qquad(n=1,2,\cdots).$$

随机变量$\xi(\omega)$的分布列为

$$p(\xi=0)=1,$$

则对$\forall\varepsilon>0$，

$$p(\omega|\,|\xi_n(\omega)-\xi(\omega)|\geqslant\varepsilon)=P(|\xi_n|\geqslant\varepsilon)$$

$$=P(\xi_n=n)=\frac{1}{n}\to 0,$$

所以　$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$。

但是　$E\xi_n^k=n^k\cdot\frac{1}{n}=n^{k-1}$，

$$E\xi^k=0^k\cdot 1=0.$$

故对任何$k\geqslant 1$，都有

$$E\xi_n^k\not\longrightarrow E\xi^k.$$

进一步的讨论

可以证明下述结果：

定理　如果$\{\xi_n\}$为正态随机变量序列，且$\xi_n\xrightarrow{P}\xi$，则有

$$E\xi_n \to E\xi, \qquad D\xi_n \to D\xi.$$

为此先证明一个引理：

引理 设 $\{F_n(x)\}$ 为一列正态分布函数，它弱收敛于分布函数 $F(x)$，则 $F(x)$ 也是正态分布函数。

证明 设 F_n、F 所对应的特征函数为 φ_n、ϕ，由假设，不妨可设

$$\phi_n(t) = \exp\left\{ia_n t - \frac{1}{2}\sigma_n^2 t^2\right\}.$$

由于 $F_n(x) \xrightarrow{w} F(x)$，

所以 $\phi(t) = \lim\limits_{n\to\infty}\phi_n(t) = \lim\limits_{n\to\infty}\exp\left\{ia_n t - \frac{1}{2}\sigma_n^2 t^2\right\}$ (*—1)。

因此如果能证明极限 $\lim\limits_{n\to\infty} a_n$ 和 $\lim\limits_{n\to\infty}\sigma_n^2$ 存在，则本引理就得以证明。由(*—1)得

$$\exp\left\{-\frac{1}{2}\sigma_n^2 t^2\right\} = |\phi_n(t)| \to |\phi(t)|.$$

因特征函数 $\phi(t)$ 连续且 $\phi(0)=1$，故在 $t=0$ 附近必存在 $t_0 \neq 0$，使 $\phi(t_0)\neq 0$，从而

$$-\frac{1}{2}\sigma_n^2 t_0^2 \longrightarrow \log|\phi(t_0)| \neq -\infty,$$

从而证得极限

$$\lim_{n\to\infty}\sigma_n^2 \triangleq \sigma^2$$

存在。

由上段证明以及假设 $F_n \xrightarrow{w} F$ 可知当 $t\in[0,1]$ 时，一致地有

$$\varphi_n(t) \to \phi(t)$$

和 $$\exp\left\{\frac{1}{2}\sigma_n^2 t^2\right\} \longrightarrow \exp\left\{\frac{1}{2}\sigma^2 t^2\right\}.$$

因此　$\exp\{ia_n t\} = \phi_n(t)\cdot\exp\left\{\frac{1}{2}\sigma_n^2 t^2\right\} \xrightarrow{\text{一致}}$

$$\phi(t)\exp\left\{\frac{1}{2}\sigma^2 t^2\right\}, \tag{*-2}$$

故得　$\left|\phi(t)\exp\left\{\frac{1}{2}\sigma^2 t^2\right\}\right| = \lim_{n\to\infty}|e^{ia_n t}| = 1$。

现取积分路线

$$C_n: \ z = \exp\{ia_n t\}, \quad 0\leqslant t\leqslant 1.$$

$$C: \ z = \phi(t)\exp\left\{\frac{1}{2}\sigma^2 t^2\right\}, \quad 0\leqslant t\leqslant 1.$$

既然在积分路线上$|z| = 1\neq 0$，再根据(*-2)得

$$ia_n = \int_{C_n}\frac{dz}{z}\longrightarrow\int_{C}\frac{dz}{z},$$

于是证得极限$\lim\limits_{n\to\infty} a_n \triangleq a$的存在性．进而有

$$\phi(t) = \exp\left\{iat - \frac{1}{2}\sigma^2 t^2\right\},$$

所以$F(x)$是数学期望为a，方差为σ^2的正态分布函数。

利用此引理就可容易地证明上述定理。事实上，若设$F_n(x)$和$F(x)$分别为ξ_n和ξ的分布函数，由$\xi_n\xrightarrow{P}\xi$推得$\xi_n\xrightarrow{L}\xi$，从而$F_n(x)\xrightarrow{w}F(x)$，再由引理知$F(x)$为正态分布函数，且$E\xi_n\longrightarrow E\xi$，$D\xi_n\longrightarrow D\xi$。

6—10　$\xi_n\xrightarrow{P}\xi$但$\xi_n\xrightarrow{a.s}\xi$不成立的例子

有关概念和命题

1　（依概率收敛）见例6—8。

2　（几乎处处收敛或概率1收敛）设$\xi(\omega)$，$\xi_1(\omega)$，$\xi_2(\omega)$，

…, $\xi_n(\omega)$, …均为随机变量，如果

$$P(\omega \mid \lim_{n\to\infty}\xi_n(\omega)=\xi(\omega))=1,$$

则称$\{\xi_n(\omega)\}$几乎处处（或概率1）收敛于$\xi(\omega)$，并记为

$$\xi_n(\omega)\xrightarrow{a,s}\xi(\omega).$$

3 若 $\xi_n(\omega)\xrightarrow{a,s}\xi(\omega)$,

则 $$\xi_n(\omega)\xrightarrow{P}\xi(\omega).$$

4 （波雷尔—康特立定理）设$A_1, A_2, A_3, \cdots$是随机事件序列

（1） 若$\sum_{n=1}^{\infty}P(A_n)<\infty$,

则 $$P(\overline{\lim_{n\to\infty}}A_n)=0$$

（上极限事件$\overline{\lim_{n\to\infty}}A_n$的概念见例1—16）。

（2） 若$\{A_n\}$是相互独立的，且

$$\sum_{n=1}^{\infty}P(A_n)=\infty,$$

则 $$P(\overline{\lim_{n\to\infty}}A_n)=1.$$

下面的例子都将表明，由$\xi_n\xrightarrow{P}\xi$不能保证$\xi_n\xrightarrow{a,s}\xi$成立。

反例

例1 取$\Omega=(0,1]$，$\mathscr{F}$为$(0,1]$中全体波雷尔集类，P为勒贝格测度，则$(\Omega,\mathscr{F},P)$为一概率空间．令

$$\eta_1^{(1)}(\omega)\equiv 1,\ \omega\in(0,1].$$

$$\eta_1^{(2)}(\omega)=\begin{cases}1, & \omega\in\left(0,\ \frac{1}{2}\right]\\ 0, & \omega\in\left(\frac{1}{2},\ 1\right]\end{cases};$$

$$\eta_2^{(2)}(\omega)=\begin{cases}0, & \omega\in\left(0,\ \frac{1}{2}\right],\\ 1, & \omega\in\left(\frac{1}{2},\ 1\right].\end{cases}$$

一般地，把$(0,1]$分成k个等长区间，而令

$$\eta_i^{(k)}(\omega)=\begin{cases}1, & \omega\in\left(\frac{i-1}{k},\ \frac{i}{k}\right]\\ 0, & \omega\overline{\in}\left(\frac{i-1}{k},\ \frac{i}{k}\right]\end{cases}\quad\left(\begin{array}{l}i=1,2,\cdots,k\\ k=1,2,\cdots,\end{array}\right)$$

定义 $\xi_1(\omega)=\eta_1^{(1)}(\omega)$，$\xi_2(\omega)=\eta_1^{(2)}(\omega)$，$\xi_3(\omega)=\eta_2^{(2)}(\omega)$，

$\xi_4(\omega)=\eta_1^{(3)}(\omega)$，$\xi_5(\omega)=\eta_2^{(3)}(\omega)$，$\cdots$

则$\{\xi_n(\omega)\}$是一列随机变量，且$\xi_n(\omega)\xrightarrow{P}0$，这是因为对$\forall 0<\varepsilon<1$，当$n\to\infty$时，有

$$P(\omega\,|\,|\xi_n(\omega)-0|\geqslant\varepsilon)$$

$$=P\left(\omega\in\left(\frac{i-1}{k_n},\ \frac{i}{k_n}\right]\right)=\frac{1}{k_n}\to 0.$$

$\left(\text{因为 } n=i+\frac{k_n(k_n-1)}{2},\ \therefore\text{当 } n\to\infty \text{ 时，} k_n\to\infty\right)$，但是在$(0,1]$中任意一点$\omega_0$，当$n\to\infty$时，

$$\xi_n(\omega_0)\to 0$$

不能成立。这是因为当$\omega_0\in(0,1]$时，对固定k必有如下形式的i，使得

$$\omega_0\in\left(\frac{i-1}{k},\ \frac{i}{k}\right],$$

从而 $\eta_i^{(k)}(\omega_0)=1$。

换言之，当我们沿数列

$$\xi_1(\omega_0),\xi_2(\omega_0),\xi_3(\omega_0),\cdots$$

看下去，不论怎样远，总有等于1的数，所以$\xi_n(\omega_0)\to 0$不能成

立，也即$\xi_n(\omega)$在（0,1］上处处不收敛于0，因而它更不能几乎处处收敛于0。

由此可见，随机变量序列依概率收敛的概念，较几乎处处收敛的概念为弱，较处处收敛的概念更弱。

例2 设随机变量$\xi_n(\omega)$的分布列定义如下：

$$P\left(\xi_n=\frac{1}{n}\right)=1-\frac{1}{n},\quad P(\xi_n=n+1)=\frac{1}{n}$$

$$(n=1,\ 2,\ 3,\ \cdots).$$

并假定$\xi_n(n=1,2,3,\cdots)$相互独立。

不难证明$\xi_n\xrightarrow{P}0$，但$\xi_n\xrightarrow{a,s}0$不难成立。事实上，若令

$$A_n=\{\omega|\xi_n(\omega)=n+1\},$$

则有 $$P(A_n)=\frac{1}{n},$$

从而 $$\sum_{n=1}^{\infty}P(A_n)=\sum_{n=1}^{\infty}\frac{1}{n}=\infty.$$

由波雷尔——康特立引理知

$$P(\overline{\lim_{n\to\infty}}A_n)=1.$$

由于 $$P(\overline{\lim_{n\to\infty}}A_n)=P(\omega|\omega\text{属于无穷多个}A_n)$$

$$=P(\omega|\text{对任何自然数}N,\text{总}\exists n>N,\text{使}\xi_n(\omega)=n+1)$$

$$\leqslant P(\omega|\xi_n(\omega)\text{不收敛于}0),$$

于是得到$P(\omega|\xi_n(\omega)\text{不收敛于}0)=1$，

因而ξ_n不可能几乎处处收敛于0。

进一步的讨论

虽然在一般情况下从随机变量的依概率收敛性推不出几乎处处收敛性，但是下列结果还是相当有用的：

1 （定理）若$\{\xi_n\}$是严格递减的正随机变量序列，则

$$\xi_n \xrightarrow{P} 0 \Longleftrightarrow \xi_n \xrightarrow{a,s} 0.$$

证明 充分性显然。

下证必要性：

由[1]p43可知

$$\xi_n \xrightarrow{a,s} 0 \Longleftrightarrow \lim_{n\to\infty} P\left\{\bigcup_{m=n}^{\infty}[|\xi_m|\geqslant\varepsilon]\right\}=0.$$

对$\forall\varepsilon>0$。由于$\xi_n\searrow 0$，∴当$m<k$时，

$$|\xi_m|>|\xi_k|,$$

故
$$\{|\xi_k|\geqslant\varepsilon\}\subset\{|\xi_m|\geqslant\varepsilon\},$$

∴
$$\left\{\bigcup_{m=n}^{\infty}[|\xi_m|\geqslant\varepsilon]\right\}=[|\xi_n|\geqslant\varepsilon],$$

$$P\left\{\bigcup_{m=n}^{\infty}[|\xi_m|\geqslant\varepsilon]\right\}=P\{|\xi_n|\geqslant\varepsilon\},$$

$$\lim_{n\to\infty} P\left\{\bigcup_{m=n}^{\infty}[|\xi_m|\geqslant\varepsilon]\right\}=\lim_{n\to\infty}P\{|\xi_n|\geqslant\varepsilon\}=0.$$

由此即知在$\xi_n\searrow 0$时由$\xi_n\xrightarrow{P}0$，可得$\xi_n\xrightarrow{a,s}0$，证毕。

2 （黎斯定理）若$\xi_n\xrightarrow{P}\xi$，则存在一个子列$\{\xi_{n_k}, k\geqslant 1\}$使得$\xi_{n_k}\xrightarrow{a,s}\xi$ $(k\to\infty)$（证明参阅[3]）。

6—11 波雷尔—康特立引理(1)的逆不成立的例子

有关概念和命题

波雷尔—康特立引理见例6—10。

对随机事件序列$\{A_n\}$放弃独立性的要求，那么即使

$P(\overline{\lim\limits_{n\to\infty}}A_n)=0$，但并不能保证$\sum\limits_{n=1}^{\infty}P(A_n)<\infty$成立。请看下面的反例。

反例

设$\Omega=[0,1]$，$\mathscr{F}$是Ω的波雷尔子集类，而P是勒贝格测度。令

$$A_n=\left(0,\ \frac{1}{n}\right)\quad(n=1,2,\cdots),$$

则 $$A_1\supset A_2\supset\cdots\supset A_n\supset\cdots,\qquad P(A_n)=\frac{1}{n}.$$

故有 $$P(\overline{\lim_{n\to\infty}}A_n)=P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\right)=P\left(\bigcap_{n=1}^{\infty}A_n\right)$$

$$=P(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}P(A_n)=\lim_{n\to\infty}\frac{1}{n}=0.$$

但是 $$\sum_{n=1}^{\infty}P(A_n)=\sum_{n=1}^{\infty}\frac{1}{n}=\infty,$$

在本例中诸A_n显然是相关的。

6—12 $\xi_n\xrightarrow{P}\xi$但$\xi_n\xrightarrow{r}\xi$不成立的例子

有关概念和命题

1 （**依概率收敛**）见例6—8。

2 （**r—阶收敛**）设随机变量ξ_n和ξ，有

$$E|\xi_n|^r<+\infty,\qquad E|\xi|^r<+\infty.$$

其中$r>0$为常数，如果

$$\lim_{n\to\infty}E|\xi_n-\xi|^r=0,$$

则称$\{\xi_n\}$ r—阶收敛于ξ，并记为 $\xi_n \xrightarrow{r} \xi$。特别当 $r=2$ 时，称为2—阶收敛或均方收敛，它是最重要的情况。

3 （定理）若$\{\xi_n\}$ r—阶收敛，则$\{\xi_n\}$依概率收敛。即由 $\xi_n \xrightarrow{r} \xi$，可得 $\xi_n \xrightarrow{P} \xi$。但反之不然。

反例

设随机变量序列$\{\xi_n\}$，其概率分布定义为

$$P(\xi_n=0)=1-\frac{1}{n^r}, \quad P(\xi_n=n)=\frac{1}{n^r}$$

$$(r>0, \ n=1,2,3,\cdots),$$

则
$$E|\xi_n|^r=n^r \cdot \frac{1}{n^r}=1,$$

于是
$$\xi_n \stackrel{r}{\not\longrightarrow} 0。$$

但是我们有
$$\xi_n \xrightarrow{P} 0。$$

事实上，对$\forall \varepsilon>0$，若$\varepsilon \leqslant n$，则

$$P(|\xi_n|>\varepsilon)=P(|\xi_n|=n)=\frac{1}{n^r}。$$

若$\varepsilon>n$，则
$$P(|\xi_n|>\varepsilon)=0,$$

因此，当$n\to\infty$时，总有
$$\lim_{n\to\infty} P(|\xi_n|>\varepsilon)=0。$$

6—13 设$0<s<r$，$\xi_n \xrightarrow{s} \xi$但$\xi_n \xrightarrow{r} \xi$不成立的例子

有关概念和定理

1 （r—阶收敛）见例6—12。

2 定理 若$r>s>0$，则由 $\xi_n \xrightarrow{r} \xi$ 可推出 $\xi_n \xrightarrow{s} \xi$。

证明 记

$$\beta_n = E|\xi|^n < \infty,$$

则由李雅普诺夫不等式(见[2]p103)知，对$\forall k$，$2 \leqslant k \leqslant n$，有

$$\beta_{k-1}^{\frac{1}{k-1}} \leqslant \beta_k^{\frac{1}{k}},$$

于是当$s<r$时，有

$$E|\xi_n-\xi|^r = \{[E|\xi_n-\xi|^s]^{\frac{1}{s}}\}^s \leqslant \{[E|\xi_n-\xi|^r]^{\frac{1}{r}}\}^s,$$

$\because \quad \xi_n \xrightarrow{r} \xi,$

$\therefore \quad E|\xi_n-\xi|^r \leqslant [E|\xi_n-\xi|^r]^{\frac{s}{r}} \to 0 \quad (n\to\infty)$

证毕。

但定理的逆不真，我们有如下的

反例

设$\{\xi_n\}$为随机变量序列，其概率分布为

$$P(\xi_n=0) = 1-\frac{1}{n^2}, \quad P(\xi_n=n) = \frac{1}{n^2}$$

$$(n=1, 2, 3, \cdots),$$

则 $\quad E|\xi_n| = n \cdot \frac{1}{n^2} = \frac{1}{n} \to 0 \quad (n\to\infty),$

即 $\quad \xi_n \xrightarrow{1} 0,$

但 $\quad E|\xi_n|^2 = n^2 \cdot \frac{1}{n^2} = 1 \not\to 0,$

因而 $\quad \xi_n \xrightarrow{2} 0$

不成立。

6—14 有关r-阶收敛与几乎处处收敛之间关系的反例

(1) $\xi_n \xrightarrow{r} \xi$，但$\xi_n \xrightarrow{a.s} \xi$不成立的例子

(2) $\xi_n \xrightarrow{a.s} \xi$，但$\xi_n \xrightarrow{r} \xi$不成立的例子

(3) $\xi_n \xrightarrow{r} \xi$，和$\xi_n \xrightarrow{a.s} \xi$都成立的例子

有关概念和命题

1 （$\{\xi_n\}r$—阶收敛）见例6—12.

2 （$\{\xi_n\}$几乎处处收敛）见例6—10.

3 （定理）$\xi_n \xrightarrow{a.s} \xi$的充要条件是，对$\forall \varepsilon > 0$，有

$$\lim_{n\to\infty} P\left(\bigcup_{j=n}^{\infty}\{|\xi_j - \xi| \geqslant \varepsilon\}\right) = 0. \qquad (*)$$

反例

(1)的反例　在例 6—10 款的反例中我们已经证明了 $\xi_n \xrightarrow{a.s} 0$不成立. 可是，对任意$r>0$,

$$E|\eta_i^{(k)}(\omega)|^r = P\left(\left(\frac{i-1}{k}, \frac{i}{k}\right]\right) = \frac{1}{k}.$$

当$n\to\infty$时，由

$$n = \frac{k(k-1)}{2} + i \leqslant \frac{k(k-1)}{2} + k.$$

知$k\to\infty$，因此

$$\lim_{n\to\infty} E|\xi_n(\omega)|^r = \lim_{n\to\infty} E|\eta_i^{(k)}(\omega)|^r$$

$$= \lim_{k\to\infty}\frac{1}{k} = 0,$$

所以 $\xi_n(\omega)\xrightarrow{r}0$。

（2）的反例 设$\Omega=[0,1]$，$\mathscr{F}$为$[0,1]$中的波雷尔集类，P为勒贝格测度.

对每个$n\geqslant1$、$r>0$，令

$$\xi_n(\omega)=\begin{cases}n, & \omega\in\left[0,\ \dfrac{1}{n^r}\right],\\ 0, & \omega\in\left(\dfrac{1}{n^r},\ 1\right].\end{cases}$$

显然，对任意 $\omega\in[0,1]$有$\lim\limits_{n\to\infty}\xi_n(\omega)=0$，故$\xi_n(\omega)\xrightarrow{a.s}0$成立。但对每个$n\geqslant1$及$r>0$，有

$$E|\xi_n|^r=n^r\cdot P\left(\left[0,\ \frac{1}{n^r}\right]\right)$$

$$=n^r\cdot\frac{1}{n^r}=1,$$

故$\xi_n\xrightarrow{r}0$ 不成立。

（3）的例子 设$\{\xi_n\}$是随机变量序列，对任意$n\geqslant1$，ξ_n的概率分布为

$$P\left(\xi_n=\frac{1}{n}\right)=\frac{1}{2},\quad P\left(\xi_n=-\frac{1}{n}\right)=\frac{1}{2},$$

则对任何$r>0$，

$$E|\xi_n|^r=\left|-\frac{1}{n}\right|^r\cdot\frac{1}{2}+\left|\frac{1}{n}\right|^r\cdot\frac{1}{2}=\frac{1}{n^r}\to0,$$

从而 $\xi_n\xrightarrow{r}0$。

另一方面，对$j<k$，$|\xi_j|>|\xi_k|$，于是

$$\{\omega\,|\,|\xi_k|>\varepsilon\}\subset\{\omega\,|\,|\xi_j|>\varepsilon\}$$

对任意$\varepsilon>0$成立，从而得到

$$\bigcup_{j=n}^{\infty}\{|\xi_j|>\varepsilon\}=\{|\xi_n|>\varepsilon\}.$$

选取$n>\frac{1}{\varepsilon}$，我们有

$$P\left[\bigcup_{j=n}^{\infty}\{|\xi_j|>\varepsilon\}\right]=P[|\xi_n|>\varepsilon]$$

$$\leqslant P\left[|\xi_n|>\frac{1}{n}\right]=0.$$

由本款有关概念和命题中的定理中的(*)式可知

$$\xi_n\xrightarrow{a,s}0.$$

(*)6—15 当$F_n\overset{C}{\nrightarrow}F$时，$\int_{-\infty}^{\infty}gdF_n\nrightarrow\int_{-\infty}^{\infty}gdF$ 的例子

有关的定理

拓广的Helly—Bray定理 设g是R_1上的有界、实值连续函数，设$\{F_n\}$是R_1上的一致、有界、非降、右连续函数列，$F_n\xrightarrow{C}F$，F是某一定义在R_1上的函数，则

$$\int_{-\infty}^{\infty}gdF_n\longrightarrow\int_{-\infty}^{\infty}gdF\quad(n\to\infty)$$

(见[1]P137定理3.1.4)。但是当$F_n\overset{C}{\nrightarrow}F$ 时，定理的结论不再成立，请看

反例.

设G是一个分布函数，因而

$$G(-\infty)=0,\qquad G(+\infty)=1.$$

定义 $F_n(x)=G(x+n)\qquad(x\in R_1)$,

则 $F_n(x)\xrightarrow{W}F\equiv1.$

但 $\quad F_n \overset{C}{\not\longrightarrow} F$，

设$g(x)\equiv 1$（当$x\in R_1$时），则g是R_1上的有界、实值、连续函数且

$$\int_{-\infty}^{\infty} g dF_n = \int_{-\infty}^{\infty} dG(x+n) = 1 \quad (n\geqslant 1).$$

而 $\int_{-\infty}^{\infty} g dF = \int_{-\infty}^{\infty} dF = 0$，

因此 $\quad \int_{-\infty}^{\infty} g dF_n \not\longrightarrow \int_{-\infty}^{\infty} g dF$

(*)6—16 ξ_n 依分布收敛到ξ（记为$\xi_n \xrightarrow{L} \xi$），但$\xi_n$ 之矩母函数Mn(S)不收敛到ξ之矩母函数M(S)的例子

有关命题

定理 记ξ_n之矩母函数为$M_n(S)$，ξ之矩母函数为$M(s)$，则当$Mn(S)\longrightarrow M(S)$时必有$\xi_n \xrightarrow{L} \xi$（见 J.H.Curtiss：Ann Math Stat13(1942)P430—433）。但定理之逆不真，我们有如下的

反例

考虑分布函数

$$F_n(x)=\begin{cases} 0, & x<-n, \\ \dfrac{1}{2}+C_n \mathrm{tg}^{-1}(nx), & -n\leqslant x<n, \\ 1, & x\geqslant n. \end{cases}$$

其中 $\quad C_n=\dfrac{1}{2\mathrm{tg}^{-1} n^2}$.

当$n\to\infty$时，

$$F_n(x) \longrightarrow F(x) = \begin{cases} 0, & x<0, \\ 1, & x \geq 0. \end{cases}$$

在$F(x)$的一切连续点成立。故

$$\xi_n \xrightarrow{L} \xi,$$

然而$F_n(x)$对应之矩母函数

$$M_n(S) = \int_{-n}^{n} C_n e^{sx} \frac{n}{1+n^2x^2} dx \quad 对一切S存在而F(x)$$

对应之矩母函数为

$$M(S) = 1 \qquad 对一切S。$$

但 $\qquad M_n(S) \not\longrightarrow M(S),$

这是因为$M_n(S) \longrightarrow \infty$（当$S \neq 0$时）。

(*)6—17 矩母函数的极限函数不是矩母函数的例子

例1 设$\{\xi_n\}$是随机变量序列，满足

$$P(\xi_n = -n) = 1, \quad n = 1, 2, \cdots$$

于是 $\qquad M_n(S) = Ee^{S\xi_n} = e^{-Sn} \to 0 \quad (n\to\infty) 当S>0时$

$$M_n(S) \longrightarrow +\infty (n\to\infty) \quad 当S<0时,$$

$$M_n(S) \longrightarrow 1 \quad (n\to\infty) \quad 当S=0时,$$

$$\therefore \quad M_n(S) \xrightarrow[(n\to\infty)]{} M(S) = \begin{cases} 0 & (S>0), \\ 1 & (S=0), \\ \infty & (S<0). \end{cases}$$

但$M(S)$并非矩母函数。事实上，若ξ_n的分布函数为$F_n(x)$，则

$$F_n(x) = \begin{cases} 0 & 当x<-n \\ 1 & 当x\geq -n \end{cases} \longrightarrow F(x) = 1 \quad 对一切x,$$

而$F(x)$不是分布函数。

第七章　与大数定律、中心极限定理有关的反例

7—1　随机变量序列$\{\xi_n\}$不服从于大数定律的例子

有关概念

1　设ξ_1，ξ_2，…，ξ_n，…是随机变量序列，

令
$$\eta_n=\frac{1}{n}(\xi_1+\xi_2+\cdots+\xi_n)。$$

如果存在一个常数序列a_1，a_2，…，a_n，…，使对$\forall\varepsilon>0$，恒有
$$\lim_{n\to\infty}P(|\eta_n-a_n|<\varepsilon)=1$$
成立，则称随机变量序列$\{\xi_n\}$服从大数定律。

2　（辛钦大数定律）设随机变量序列$\{\xi_k\}$独立同分布，则$\{\xi_k\}$服从大数定律的充要条件是ξ_1有有穷的数学期望。

下面举几个不服从大数定律的随机变量序列的例子。

反例

例1　设随机变量ξ_k的分布定义如下：
$$P(\xi_k=2^k)=\frac{1}{2},\ P(\xi_k=-2^k)=\frac{1}{2}.$$
$$(k=1,\ 2,\ 3,\ \cdots)。$$

并设随机变量序列$\{\xi_k\}$相互独立，因而

$$P(\xi_{n-1}=2^{n-1},\ \xi_n=2^n)$$
$$=P(\xi_{n-1}=2^{n-1})\cdot P\ (\xi_n=2^n)=\frac{1}{4}.$$

当$\xi_{n-1}=2^{n-1}$，$\xi^n=2^n$时，有

$$|\xi_1+\xi_2+\cdots+\xi_n|\geqslant|\xi_n+\xi_{n-1}|-|\xi_1+\cdots+\xi_{n-2}|$$
$$\geqslant(2^n+2^{n-1})-(2+2^2+\cdots+2^{n-2})>2^n.$$

所以当n充分大时，有

$$\frac{1}{n}|\xi_1+\cdots+\xi_n|\geqslant\frac{2^n}{n}\geqslant 1.$$

因此，事件

$$(\xi_{n-1}=2^{n-1},\ \xi_n=2^n)\subset\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i\right|\geqslant\frac{2^n}{n}\right)$$

于是，当n充分大时，对$\varepsilon=1$，

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i\right|\geqslant\varepsilon\right)\geqslant P(\xi_{n-1}=2^{n-1},\ \xi_n=2^n)$$
$$=\frac{1}{4}\not\longrightarrow 0.$$

从而表明随机变量序列$\{\xi_k\}$不服从大数定律。

例2 设骰子A有四面红和两面白；而骰子B有两面红四面白。抛一个钱币，若出现正面，接着就连续掷骰子A；若出现反面，就掷骰子B。令随机变量

$$\xi_k=\begin{cases}1, & \text{第}k\text{次掷出红面},\\ 0, & \text{其他}\end{cases}$$

则大数定律不适用于$\{\xi_k\}$。

我们首先证明$\{\xi_k\}$同分布，但不独立。事实上

$$P(\xi_k=1)=P\ (\text{第}k\text{次掷出红面})$$

$$=P(\text{抛钱币出正面})\cdot P(\text{第}k\text{次出红面}\mid\text{掷骰子}A)$$
$$+P(\text{抛钱币出反面})P(\text{第}k\text{次出红面}\mid\text{掷骰子}B)$$
$$=\frac{1}{2}\cdot\frac{2}{3}+\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{2}.$$

类似地可得

$$P(\xi_k=0)=\frac{1}{2}\qquad(k=1,\ 2,\ 3,\ \cdots),$$

从而 $\quad E(\xi_k)=\dfrac{1}{2},\quad E\left(\sum_{k=1}^{n}\xi_k\right)=\dfrac{n}{2}.$

因为在A（指掷骰子A事件）发生的条件下，$\{\xi_k\}$相互独立，同分布，其分布列为

$$P(\xi_k=1\mid A)=\frac{2}{3},\quad P(\xi_k=0\mid A)=\frac{1}{3},$$

且 $\quad E(\xi_k\mid A)=\dfrac{2}{3}.$

同理，在B（指掷骰子B事件）发生的条件下，$\{\xi_k\}$也相同独立，同分布，而且

$$P(\xi_k=1\mid B)=\frac{1}{3},\quad P(\xi_k=0\mid B)=\frac{2}{3},\quad E(\xi_k\mid B)=\frac{1}{3},$$

因而 $\quad P(\xi_{k-1}=1,\ \xi_k=1)=P(A)P(\xi_{k-1}=1,\ \xi_k=1\mid A)$
$$+P(B)P(\xi_{k-1}=1,\ \xi_k=1\mid B)$$
$$=\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{2}{3}+\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{3}=\frac{5}{18},$$

但 $\quad P(\xi_{k-1}=1)=P(\xi_k=1)=\dfrac{1}{2},$

所以 $\quad P(\xi_{k-1}=1,\ \xi_k=1)\neq P(\xi_{k-1}=1)\cdot P(\xi_k=1),$

因此$\{\xi_k\}$不独立。

其次证明，对$\varepsilon>0$，

$$P\left(\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-E\left(\sum_{k=1}^{n}\xi_k\right)\right|>\varepsilon\right)\not\to 0.$$

由辛钦大数定律可得

$$P\left\{\left[\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{2}{3}n\right|>\varepsilon\right]\middle|A\right\}\longrightarrow 0$$

和
$$P\left\{\left[\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{3}\right|>\varepsilon\right]\middle|B\right\}\longrightarrow 0,$$

因此
$$P\left\{\left[\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\right]\middle|A\right\}\not\to 0,$$

$$P\left\{\left[\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\right]\middle|B\right\}\not\to 0,$$

而
$$P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\right\}$$

$$=P(A)\cdot P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\middle|A\right\}$$

$$+P(B)\cdot P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\middle|B\right\}$$

$$=\frac{1}{2}\cdot P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\middle|A\right\}$$

$$+\frac{1}{2}\cdot P\left\{\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k-\frac{n}{2}\right|>\varepsilon\middle|B\right\}$$

$$\not\to 0.$$

即$\{\xi_k\}$不满足大数定律。

例3 设$\{\xi_k\}$为随机变量序列，

$$S_n=\sum_{k=1}^{n}\xi_k.$$

若$|S_n|<Cn$，且$D(Sn)>\alpha n^2$（c、α均为大于零的常数，则大数定律不能应用于$\{\xi_k\}$。

用反证法，设$\{\xi_k\}$服从大数定律，即任给$\varepsilon>0$，都有

$$\lim_{n\to\infty}P\left\{\left|\sum_{k=1}^{n}\xi_k-E\left(\sum_{k=1}^{n}\xi_k\right)\right|<n\varepsilon\right\}=1,$$

亦即 $$\lim_{n\to\infty}P\{|S_n-ES_n|<n\varepsilon\}=1. \qquad (*)$$

由假设条件知，$D(S_n)>\alpha n^2$。另一方面，若设S_n的分布函数为$F_{S_n}(x)$，则

$$\begin{aligned}D(S_n)&=\int_{-\infty}^{\infty}(x-E(S_n))^2dF_{S_n}(x)\\&=\int_{|x-E(Sn)|<n\varepsilon}(x-E(S_n))^2dF_{S_n}(x)\\&\quad+\int_{|x-E(Sn)|\geqslant n\varepsilon}(x-E(S_n))^2dF_{S_n}(x)\\&\leqslant n^2\varepsilon^2\int_{|x-E(Sn)|<n\varepsilon}dF_{S_n}(x)\\&\quad+\int_{|x-E(Sn)|\geqslant n\varepsilon}[x^2-2xE(S_n)+E^2(S_n)]\\&\quad\cdot dF_{S_n}(x),\end{aligned}$$

因为$|S_n|<cn$，从而上式右边第二个积分化为

$$\int_{|x-E(S_n)|\geqslant n\varepsilon}[x^2-2xE(S_n)+E^2(S_n)]\,dF_{S_n}(x)$$

$$\leqslant[n^2c^2+2cnE(S_n)+E^2(S_n)]\int_{|x-E(S_n)|\geqslant n\varepsilon}dF_{S_n}(x)$$
$$=[n^2c^2+2cnE(S_n)+E^2(S_n)]P(|S_n-E(S_n)|\geqslant n\varepsilon),$$

故有 $D(S_n)\leqslant n^2\varepsilon^2+[nc+E(S_n)]^2P(|S_n-E(S_n)|\geqslant n\varepsilon)$。

因为ε可任意小，再由(*)知

$$P(|S_n-E(S_n)|\geqslant n\varepsilon)$$

也可以任意小，这就要与$D(S_n)>\alpha n^2$相矛盾，因此$\{\xi_k\}$不服从大数定律。

例4 在Palya罐子模型（罐子中有b个黑球，r个红球，随机地取出一个球，然后再放回罐中，并且还多放c个同色球）中，令

$$\xi_k=\begin{cases}1 & 第k次取出黑球,\\ 0 & 第k次取出红球,\end{cases}$$

则S_n是n次取球中所取出的黑球数。此时大数定律不能应用于$\{\xi_k\}$。

事实上，不难得到

$$E(S_n)=\frac{nb}{b+r},$$

$$D(S_n)=\frac{nbr(b+r+nc)}{(b+r)^2(b+r+c)}>\frac{brcn^2}{(b+r)^2(b+r+c)}.$$

显然 $|S_n|\leqslant n<2n$，

故由例3的结论知$\{\xi_k\}$不服从大数定律。

7—2 随机变量序列$\{\xi_k\}$不满足马尔科夫条件，但服从大数定律的例子

有关概念和命题

1 （格涅坚科定理）见[18]P210—212。

对随机变量序列$\{\xi_k\}$，若记

$$\eta_n \triangleq \frac{1}{n}\sum_{k=1}^{n}\xi_k, \qquad a_n \triangleq \frac{1}{n}E\left(\sum_{k=1}^{n}\xi_k\right),$$

则$\{\xi_k\}$服从大数定律的充要条件是

$$\lim_{n\to\infty}E\left[\frac{(\eta_n-a_n)^2}{1+(\eta_n-a_n)^2}\right]=0。\qquad (*-1)$$

如果随机变量ξ_1，ξ_2，…，ξ_n，…彼此独立，则条件(*—1)等价于

$$\lim_{n\to\infty}\sum_{k=1}^{n}E\left[\frac{(\xi_k-E\xi_k)^2}{n^2+(\xi_k-E\xi_k)^2}\right]=0。\qquad (*-2)$$

2 （马尔科夫条件和马尔科夫大数定律）

对随机变量序列$\{\xi_k\}$，若

$$\frac{1}{n^2}D\left(\sum_{k=1}^{n}\xi_k\right)\longrightarrow 0, \qquad (*-3)$$

则对任给$\varepsilon>0$，均有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-\frac{1}{n}\sum_{k=1}^{n}E\xi_k\right|<\varepsilon\right\}=1$$

成立。通常称条件(*—3)为马尔科夫条件。

马尔科夫大数定律的重要意义在于，它对$\{\xi_n\}$不需要任何关于独立性的假定。

马尔科夫条件是服从大数定律的充分条件。我们可以举出不满足马尔科夫条件，但仍然服从大数定律的随机变量序列的例子。

反例

设$\{\xi_k\}$是相互独立的随机变量序列，且

$$P(\xi_k=\pm 1)=\frac{1}{2}(1-2^{-k}),\quad P(\xi_k=\pm 2^k)=2^{-k-1},$$

则 $$E\xi_k=0,$$
$$D\xi_k=E\xi_k^2=(1-2^k)+2^{2k}\cdot 2^{-k}=2^k+1+2^{-k}.$$

所以 $$D\left(\sum_{k=1}^{n}\xi_k\right)=\sum_{k=1}^{n}D\xi_k$$
$$=\sum_{k=1}^{n}(2^k+1-2^{-k})$$
$$=2^{n+1}-2+n+1-\left(\frac{1}{2}\right)^n$$
$$=2^{n+1}+n-1-\left(\frac{1}{2}\right)^n.$$

于由$\frac{2^{n+1}}{n^2}$在$n\to\infty$时不趋于0，所以
$$\frac{1}{n^2}D\left(\sum_{k=1}^{n}\xi_k\right)\not\longrightarrow 0,$$
故马尔科夫条件不满足，但是$\{\xi_k\}$仍服从大数定律．事实上，因为$\{\xi_k\}$是相互独立随机变量序列，且 $E\xi_k=0(k=1, 2, 3, \cdots)$，所以
$$\sum_{k=1}^{n}E\left[\frac{(\xi_k-E\xi_k)^2}{n^2+(\xi_k-E\xi_k)^2}\right]=\sum_{k=1}^{n}E\left(\frac{\xi_k^2}{n^2+\xi_k^2}\right)$$
$$=\sum_{k=1}^{n}\left[\left(\frac{1}{n^2+1}\right)\left(1-\frac{1}{2^k}\right)+\frac{2^{2k}}{n^2+2^{2k}}\cdot\frac{1}{2^k}\right]$$
$$=\sum_{k=1}^{n}\frac{1-\frac{1}{2^k}}{n^2+1}\quad\sum\frac{2^k}{n^2+2^{2k}}.$$

注意到 $$\sum_{k=1}^{n}\frac{\left(1-\frac{1}{2^k}\right)}{n^2+1}<\sum_{k=1}^{n}\frac{1}{n^2+1}=\frac{n}{n^2+1}\longrightarrow 0\ (n\to\infty),$$

而 $$\int_0^n \frac{2^x}{n^2+2^{2x}}dx=\frac{1}{\ln 2}\int_0^n \frac{d(2^x)}{n^2+(2^x)^2}$$

$$=\frac{1}{n\ln 2}\mathrm{tg}^{-1}\left(\frac{2^x}{n}\right)\Big|_0^n$$

$$=\frac{1}{n\ln 2}\left[\mathrm{tg}^{-1}\left(\frac{2^n}{n}\right)-\mathrm{tg}^{-1}\left(\frac{1}{n}\right)\right]\longrightarrow 0 \quad (n\to\infty).$$

由级数收敛的积分判别法可知

$$\sum_{k=1}^{n}\frac{2^k}{n^2+2^{2k}}\longrightarrow 0,$$

于是 $$\sum_{k=1}^{n}E\left[\frac{(\xi_k-E\xi_k)^2}{n^2+(\xi_k-E\xi_k)^2}\right]\longrightarrow 0.$$

由格涅坚科定理知$\{\xi_k\}$服从大数定律.

7—3 独立随机变量序列$\{\xi_k\}$不满足车贝谢夫大数定律的条件，但满足马尔科夫大数定律条件的例子

有关概念和命题

1 车贝谢夫大数定律，设ξ_1，ξ_2，…，ξ_n，…是两两不相关的随机变量序列，每一随机变量都有有限方差，并且它们一致有界，即

$$D\xi_k\leqslant C \qquad (k=1,\ 2,\ \cdots).$$

其中C为某一常数，则对任给$\varepsilon>0$，都有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-\frac{1}{n}\sum_{k=1}^{n}E\xi_k\right|<\varepsilon\right\}=1.$$

2 马尔科夫条件和马尔科夫大数定律见例7—2.

反例

设$\{\xi_k\}$是独立随机变量序列，且

$$P\{\xi_k=\pm\sqrt{\log k}\}=\frac{1}{2},$$

则$\{\xi_k\}$不满足车贝谢夫大数定律的条件。这是因为

$$E\xi_k=0,\quad D\xi_k=E\xi_k^2=\log k.$$

显然 $D\xi_k=\log k\quad(k=1,2,3,\cdots)$。

非一致有界，所以$\{\xi_k\}$不满足车贝谢夫大数定律的条件，但是$\{\xi_k\}$满足马尔科夫条件，从而它服从马尔科夫大数定律。事实上

$$D\left(\sum_{k=1}^{n}\xi_k\right)=\sum_{k=1}^{n}D\xi_k=\sum_{k=1}^{n}\log k\leqslant n\log n,$$

所以 $$\frac{1}{n^2}D\left(\sum_{k=1}^{n}\xi_k\right)\leqslant\frac{\log n}{n}\longrightarrow 0.$$

众所周知，车贝谢夫大数定律可由马尔科夫大数定律推出。本反例则表明，马尔科夫条件要比车贝谢夫大数定律的条件来得弱。

7—4 独立随机变量序列$\{\xi_k\}$不满足格涅坚科大数定律的充要条件的例子

有关概念和命题

1 格涅坚科大数定律见例7—2。

2 独立随机变量之和的特征函数等于它们特征函数的乘积。

反例

设随机变量序列ξ_1，ξ_2，…，ξ_n，…相互独立，均服从哥西分布，密度函数为

$$f(x)=\frac{1}{\pi}\frac{1}{1+x^2}$$

则它们不满足格涅坚科大数定律的充要条件．证明如下：

因为诸$\xi_k(k=1, 2, 3, \cdots)$的密度函数

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}.$$

其特征函数为$e^{-|t|}$．利用独立性得$\sum\limits_{k=1}^{n}\xi_k$的特征函数为

$$(e^{-|t|})^n=e^{-n|t|}$$

再根据分布函数与特征函数一一对应，由此得$\sum\limits_{k=1}^{n}\xi_k$的密度函数为$\frac{1}{\pi}\cdot\frac{n}{n^2+x^2}$．从而

$$E\frac{\left(\sum\limits_{k=1}^{n}\xi_k\right)^2}{n^2+\left(\sum\limits_{k=1}^{n}\xi_k\right)^2}=\int_{-\infty}^{\infty}\frac{t^2}{n^2+t^2}\cdot\frac{1}{\pi}\frac{n}{n^2+t^2}dt$$

$$=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{nt^2}{(n^2+t^2)^2}dt=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x^2}{(1+x^2)^2}dx$$

$$=\frac{1}{\pi}\int_{-\infty}^{\infty}\left[\frac{1}{(1+x^2)}-\frac{1}{(1+x^2)^2}\right]dx$$

$$=\frac{1}{\pi}\left[\operatorname{arctg}x-\frac{x}{2(1+x^2)}-\frac{1}{2}\operatorname{arctg}x\right]_{-\infty}^{\infty}$$

$$=\frac{1}{\pi}\left(\frac{\pi}{4}+\frac{\pi}{4}\right)=\frac{1}{2}\not\to 0.$$

7—5 马尔科夫条件满足，但柯尔莫哥洛夫强大数条件不满足的例子

有关概念和命题

1 （马尔科夫条件）见例7—2。

2 （柯尔莫哥洛夫强大数定律）设$\{\xi_k\}$是独立随机变量序列，且

$$\sum_{k=1}^{\infty}\frac{D\xi_k}{k^2}<\infty,$$

则
$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(\xi_i-E\xi_i)=0\right)=1.$$

反例

设$\{\xi_k\}$是独立随机变量序列、方差

$$D\xi_k=\frac{k+1}{\log(k+1)}\qquad(k=1,\ 2,\ \cdots),$$

则马尔科夫条件对$\{\xi_k\}$满足．这是因为，$\{\xi_k\}$独立，所以

$$\lim_{n\to\infty}\frac{1}{n^2}D\left(\sum_{k=1}^{n}\xi_k\right)=\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}D\xi_k$$

$$=\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\frac{k+1}{\log(k+1)}=\lim_{n\to\infty}\frac{\sum\limits_{k=1}^{n}\dfrac{k+1}{\log(k+1)}}{n^2}$$

令$y_n=n^2$，$x_n=\sum\limits_{k=1}^{n}\dfrac{k+1}{\log(k+1)}$，

则因为 $y_{n+1}>y_n$,

且 $$\lim_{n\to\infty} y_n = +\infty.$$

以及 $$\lim_{n\to\infty}\frac{x_{n+1}-x_n}{y_{n+1}-y_n}=\lim_{n\to\infty}\frac{\left(\dfrac{n+2}{\log(n+2)}\right)}{(n+1)^2-n^2}$$

$$=\lim_{n\to\infty}\frac{\left(\dfrac{1}{\log(n+1)}\right)}{\left(\dfrac{2n+1}{n+2}\right)}=0.$$

所以 $$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}\dfrac{k+1}{\log(k+1)}}{n^2}=\lim_{n\to\infty}\frac{x_n}{y_n}$$

$$=\lim_{n\to\infty}\frac{x_{n+1}-x_n}{y_{n+1}-y_n}=0.$$

可见马尔科夫条件是满足的. 但是柯尔莫哥洛夫强大数定律的条件不满足. 因为条件要求 $\sum_{k=1}^{\infty}\frac{D\xi_k}{k^2}$ 收敛，即要考察数列

$$a_n \triangleq \sum_{k=1}^{n}\frac{k+1}{k^2\log(k+1)}$$

的收敛性. 因为

$$a_n \geqslant \sum_{k=1}^{n}\frac{k+1}{(k+1)^2\log(k+1)}=\sum_{k=1}^{n}\frac{1}{(k+1)\log(k+1)}$$

$$=\sum_{k=2}^{n}\frac{1}{k\log k}\triangleq b_n.$$

由于广义积分 $\int_2^{\infty}\frac{dx}{x\log x}$ 发散，所以数列 $\{b_n\}$ 发散，从而数列 $\{a_n\}$ 发散，故柯尔莫哥洛夫强大数定律条件不满足.

进一步的讨论

设 $\{\xi_n\}$ 是一组相互独立具有有限方差的随机变量序列，若

$$\sum_{n=1}^{\infty}\frac{D\xi_n}{n^2}<\infty,$$

则必有 $\lim\limits_{n\to\infty}\frac{1}{n^2}\sum\limits_{k=1}^{n}D\xi_k=0$。

也就是说，在这种场合下，如果柯尔莫哥洛夫强大数定律的条件满足，则马尔科夫大数定律的条件必定满足。证明如下：

因为 $\sum\limits_{n=1}^{\infty}\frac{D\xi_n}{n^2}<\infty$，

故对任给$\varepsilon>0$，$\exists M$使得

$$\sum_{k=M+1}^{\infty}\frac{D\xi_k}{k^2}<\frac{\varepsilon}{2}$$

而对任给$n>M$。

$$\frac{1}{n^2}\sum_{k=1}^{n}D\xi_k=\frac{1}{n^2}\sum_{k=1}^{M}D\xi_k+\frac{1}{n^2}\sum_{k=M+1}^{n}D\xi_k$$

$$\leqslant\frac{1}{n^2}\sum_{k=1}^{M}D\xi_k+\sum_{k=M+1}^{\infty}\frac{D\xi_k}{k^2}.$$

因为方差$D\xi_k$为有限数，所以$\sum\limits_{k=1}^{M}D\xi_k$为有限量$S_M$，故$\frac{1}{n^2}\sum\limits_{k=1}^{M}D\xi_k$ $=\frac{1}{n^2}S_M$。

取$n>S_M$，$\frac{S_M}{n^2}<\frac{1}{n}$。

故当$n>\max\left(S_M,\ M,\ \frac{2}{\varepsilon}\right)$时，便有

$$\frac{1}{n^2}\sum_{k=1}^{M}D\xi_k<\frac{\varepsilon}{2}.$$

而 $\sum_{k=M+1}^{\infty} \frac{D\xi_k}{k^2}$,

已如上述，于是

$$\frac{1}{n^2}\sum_{k=1}^{n} D\xi_k < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

因而得到$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n} D\xi_k = 0$。

7—6 独立随机变量序列$\{\xi_k\}$不满足强大数定律的例子

有关概念和命题

1 设$\{\xi_k\}$是独立随机变量序列，若

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(\xi_k - E\xi_k) = 0\right) = 1$$

则称它服从强大数定律。

2 （柯尔莫哥洛夫强大数定律）见例7—5。

反例

设独立随机变量序列$\{\xi_k\}$，其分布是

$$P(\xi_k = \pm 2^k) = \frac{1}{2} \qquad (k = 1, 2, \cdots),$$

则 $$E\xi_k = 0, \ D\xi_k = (2^k)^2 \cdot \frac{1}{2} + (-2^k)^2 \cdot \frac{1}{2} = 4^k.$$

下面用反证法证其不满足强大数定律。

若强大数定律成立，则

$$\frac{1}{n}(\xi_1 + \xi_2 + \cdots + \xi_n) \xrightarrow{a.S} 0. \tag{1}$$

因而亦应有

$$\frac{1}{n}\sum_{k=1}^{n-1}\xi_k=\frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{k=1}^{n-1}\xi_k\xrightarrow{a,S}0. \tag{2}$$

(1)－(2)得$\frac{1}{n}\xi_n\xrightarrow{a,S}0$。 (3)

但从ξ_n的取值可知，总有

$$\left|\frac{\xi_n}{n}\right|=\frac{2^n}{n}>1.$$

这显然与(3)矛盾。所以对$\{\xi_k\}$强大数定律不满足。

进一步的讨论

柯尔莫哥洛夫强大数定律提供了相互独立的随机变量序列$\{\xi_k\}$服从强大数定律的一个充分条件。下面的定理则提供了相互独立的随机变量序列服从强大数定律的必要条件。

定理 设$\{\xi_k\}$相互独立，且$E\xi_k=0$，若$\{\xi_k\}$服从强大数定律，则对任给$\varepsilon>0$有

$$\sum_{n=1}^{\infty}P\{|\xi_n|\geqslant n\varepsilon\}<\infty.$$

为了证明这一结论，首先证明下面的

引理 设$\{\xi_k\}$相互独立，且

$$P(\lim_{n\to\infty}\xi_n=0)=1.$$

则对任给$\varepsilon>0$有

$$\sum_{n=1}^{\infty}P\{|\xi_n|\geqslant\varepsilon\}<\infty.$$

事实上，由

$$P(\lim_{n\to\infty}\xi_n=0)=1$$

可知，事件$\{|\xi_n|\geqslant\varepsilon\}$只发生有限多次的概率为1．由这些事件的独立性，由波雷尔—康特立引理知

$$\sum_{n=1}^{\infty}P\{|\xi_n|\geqslant\varepsilon\}<\infty.$$

利用这个引理，就不难证明上述定理．因为$\{\xi_k\}$独立，$E\xi_k=0$，如果$\{\xi_k\}$服从强大数定律，则$\frac{S_n}{n}\xrightarrow{a.s}0$（此处$S_n \triangleq \xi_1+\xi_2+\cdots+\xi_n$）

和
$$\frac{S_{n-1}}{n-1}\xrightarrow{a.s}0,$$

因此
$$\frac{S_{n-1}}{n}=\frac{n-1}{n}\cdot\frac{S_{n-1}}{n-1}\xrightarrow{a.s}0,$$

从而
$$\frac{\xi_n}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n}\xrightarrow{a.s}0,$$

即
$$P\left(\lim_{n\to\infty}\frac{\xi_n}{n}=0\right)=1.$$

由上述引理知对任给$\varepsilon>0$有

$$\sum_{n=1}^{\infty}P\{|\xi_n|\geqslant n\varepsilon\}<\infty.$$

如果用上述定理中的必要条件来验证本款反例中的独立随机变量序列$\{\xi_k\}$，那么就很容易得到它不服从强大数定律的结论(请读者自行补证)．

利用上述定理，我们还能得到进一步的结果：

定理 如果$\sum_{k=1}^{\infty}\frac{\sigma_k^2}{k^2}$发散，则存在方差为$D\xi_k=\sigma_k^2$的独立随机变量序列$\{\xi_k\}$，使得对于这个序列强大数定律不成立(见〔13〕

第十章问题18)。

证明留给读者。

7—7 林德贝格条件不满足，但中心极限定理仍成立的例子

有关概念和命题

设$\{\xi_k\}$为相互独立随机变量序列，它们具有有限的数学期望和方差：

$$a_k = E\xi_k,\ \ b_k^2 = D\xi_k\ \ (k=1,\ 2,\ \cdots),$$

记

$$B_n^2 = \sum_{k=1}^{n} b_k^2.$$

1 （林德贝格(Linderberg)定理)若对任意$\tau>0$，有

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2 dF_k(x) = 0 \tag{1}$$

(此条件称为林德贝格条件)，则

$$\lim_{n\to\infty} P\left\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

2 （费勒(Feller)条件）

$$\lim_{n\to\infty}\max_{k\le n}\frac{b_k}{B_k} = 0 \tag{2}$$

称为费勒条件

3 定理 费勒条件(2) 等价于

$$\lim_{n\to\infty} B_n = +\infty,\ \ \lim_{n\to\infty}\frac{b_n}{B_n} = 0. \tag{3}$$

反例

设$\{\xi_k\}$是独立随机变量序列，且$\xi_k \sim N(0, 2^{-k})$，则$\{\xi_k\}$不满足林德贝格条件，但对$\{\xi_k\}$中心极限定理仍成立。事实上，由于

$$\sum_{k=1}^{\infty} b_k^2 = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1,$$

所以 $$B_n^2 = \sum_{k=1}^{n} b_k^2 \longrightarrow 1。$$

即$B_n \uparrow 1$，所以存在常数B，使$B_n \leqslant B$。注意到$a_k = 0$（$k = 1, 2, \cdots$），因

$$b_1^2 = \int_{-\infty}^{\infty} (x - a_1)^2 dF_1(x) = \int_{-\infty}^{\infty} x^2 dF_1(x)。$$

故可取$\tau > 0$，使

$$\int_{|x| > \tau B} x^2 dF_1(x) > \frac{b_1^2}{2},$$

因此 $$\lim_{n \to \infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int_{|x - a_k| > \tau B_n} (x - a_k)^2 dF_k(x)$$

$$\geqslant \frac{1}{B} \int_{|x| > \tau B} x^2 dF_1(x) > \frac{b_1^2}{2B} > 0。$$

因此林德贝格条件(1) 不满足。

但是根据正态分布的再生性知$\dfrac{\sum_{k=1}^{n} \xi_k}{B_n}$服从正态分布$N(a_n, \sigma_n^2)$，其中

$$a_n = E\left(\frac{\sum_{k=1}^{n} \xi_k}{B_n}\right) = \frac{1}{B_n} \sum_{k=1}^{n} E\xi_k = 0。$$

$$\sigma_n^2 = D\left(\frac{\sum_{k=1}^{n}\xi_k}{B_n}\right) = \frac{1}{B_n^2}\sum_{k=1}^{n} D\xi_k = 1.$$

即对任意n，

$$\frac{\sum_{k=1}^{n}\xi_k}{B_n} \sim N(0,\ 1),$$

故
$$P\left(\frac{\sum_{k=1}^{n}\xi_k}{B_n} < x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt,$$

所以中心极限定理成立。

进一步的讨论

林德贝格条件是中心极限定理成立的充分条件，但并非必要条件。本款反例就证明了这一点。不过，假如费勒条件得到满足，则林德贝格条件就成为中心极限定理成立的充分必要条件了。这就是著名的林德贝格——费勒定理（参阅〔5〕，P168）。

7—8　费勒条件不满足，但中心极限定理仍成立的例子

有关概念和定理

1　费勒条件见例7—7。

2　若$\xi\sim$均匀分布$U[a,\ b]$，则

$$E\xi = \frac{a+b}{2},\qquad D\xi = \frac{1}{12}(b-a)^2.$$

3 正态分布$N(a, \sigma^2)$的特征函数为$e^{iat-\frac{1}{2}\sigma^2t^2}$，均匀分布$U[-1, 1]$的特征函数为$\frac{\sin t}{t}$。

4 设$\eta = a\xi + b$，其中a、b为常数，则η的特征函数
$\phi_\eta(t) = e^{ibt}\phi_\xi(at)$。

5 （逆极限定理）设特征函数列$\{\varphi_n(t)\}$收敛于某一函数$\varphi(t)$，且$\varphi(t)$在$t=0$连续，则相应的分布函数列$\{F_n(x)\}$弱收敛于某一分布函数$F(x)$，而且$\varphi(t)$是$F(x)$的特征函数。

反例

设$\{\xi_k\}$是独立随机变量序列，ξ_1服从$[-1, 1]$上的均匀分布。对$k=2, 3, \cdots$，ξ_k服从$N(0, 2^{k-1})$，则对$\{\xi_k\}$成立中心极限定理，但不满足费勒条件。证明如下：

因为 $$B_n^2 = D\xi_1 + \sum_{k=2}^{n} D\xi_k = \frac{1}{3} + \sum_{k=2}^{n} 2^{k-1}$$

$$= \frac{1}{3} + \frac{2(1-2^{n-1})}{1-2} = 2^n - \frac{5}{3} \longrightarrow \infty.$$

但是 $$\frac{b_n^2}{B_n^2} = \frac{2^{n-1}}{2^n - \frac{5}{3}} \longrightarrow \frac{1}{2} \neq 0,$$

所以费勒条件不满足。$\sum_{k=1}^{n} \xi_k$的特征函数为

$$\phi(t) = \prod_{k=1}^{n} \phi_{\xi_k}(t) = \left(\frac{\sin t}{t}\right)\exp\left\{-\frac{1}{2}(2^n-2)t^2\right\}.$$

由此得到$\eta_n = \frac{1}{B_n}\sum_{k=1}^{n} \xi_k = \frac{1}{\sqrt{2^n - \frac{5}{3}}}\sum_{k=1}^{n}\xi_k$

的特征函数为

$$\psi_n(t)=\left(\sin\frac{t}{\sqrt{2^n-\frac{5}{3}}}\right)^{\frac{\sqrt{2^n-5/3}}{t}}\exp\left\{-\frac{1}{2}\frac{(2^n-2)t^2}{(2^n-5/3)}\right\}$$

$$\longrightarrow e^{-\frac{1}{2}t^2}\qquad(n\longrightarrow\infty).$$

由逆极限定理知$\eta_n\xrightarrow{L}N(0,\ 1)$，所以中心极限定理成立。

7—9 有关大数定律和中心极限定理之间关系的反例

（1）随机变量序列不服从大数定律，也不服从中心极限定理的例子；

（2）随机变量序列服从大数定律，但不服从中心极限定理的例子；

（3）随机变量序列服从中心极限定理，但不服从大数定理的例子

有关概念和命题

1 随机变量序列$\{\xi_n\}$服从大数定律的概念见例7—1。

2 （中心极限定理）设$\{\xi_n\}$为相互独立的随机变量序列，具有有限的数学期望和方差

$$E\xi_k=a_k,\quad D\xi_k=\sigma_k^2\quad(k=1,\ 2,\ \cdots)$$

令
$$B_n^2=\sum_{k=1}^{n}D\xi_k$$

$$\eta_n=\sum_{k=1}^{n}\frac{\xi_k-a_k}{B_n}\qquad(n=1,\ 2,\ \cdots).$$

若对$x\in R_1$一致地有

$$\lim_{n\to\infty}P(\eta_n<x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{1}{2}t^2}dt.$$

则称随机变量序列$\{\xi_n\}$服从中心极限定理。

3 （格涅坚科定理）见例7—2。

4 （定理＊）若ξ_1，ξ_2，…，ξ_n，…是独立随机变量序列，∃常数k_n，使

$$\max_{1\leqslant j\leqslant n}|\xi_j|\leqslant k_n\quad(n=1,\ 2,\ \cdots),$$

且
$$\lim_{n\to\infty}\frac{k_n}{B_n}=0.$$

则$\{\xi_n\}$服从中心极限定理（见〔6〕、P305）。

反例

（1）的例子　设$\{\xi_k\}$相互独立，ξ_k的分布定义如下

$$P(\xi_k=2^k)=\frac{1}{2},\quad P(\xi_k=-2^k)=\frac{1}{2}$$

$$(k=1,\ 2,\ \cdots).$$

由例7—1的反例1已知$\{\xi_k\}$不满足大数定律，下面来证明中心极限定理也不满足。事实上，

因为　$E\xi_k=0$，$D\xi_k=2^{2k}$.

所以
$$B_n^2=\sum_{k=1}^{n}D\xi_k=\sum_{k=1}^{n}2^{2k}=\sum_{k=1}^{n}4^k=\frac{4}{3}(2^{2n}-1).$$

故当n充分大时，$B_n^2>2^{2n}$，但是

$$\left|\sum_{k=1}^{n}\xi_k\right|\leqslant 2+2^2+\cdots+2^n=2(2^n-1)<2^{n+1},$$

因此，当n充分大时有

$$\left|\frac{1}{B_n^2}\sum_{k=1}^{n}\xi_k\right|<2.$$

故 $$\lim_{n\to\infty} P\left(\left|\frac{1}{B_n}\sum_{k=1}^{n}\xi_k\right|<2\right)=1\neq\Phi(2)-\Phi(-2)<1$$ 此表明$\{\xi_k\}$不可能服从中心极限定理。

（2）的例子　设随机变量ξ_k的分布定义如下：

$$P(\xi_k=\pm 2^k)=2^{-(2k+1)},\quad P(\xi_k=0)=1-2^{-2k},$$

并设随机变量序列$\{\xi_k\}$相互独立。

大数定律满足，这是因为

$$E\xi_k=0,\quad D\xi_k=2\times 2^{2k}\cdot 2^{-(2k+1)}=1,$$

故 $$B_n^2=\sum_{k=1}^{n}D\xi_k=n.$$

$$\frac{1}{n^2}\sum_{k=1}^{n}D\xi_k=\frac{n}{n^2}\longrightarrow 0.$$

由马尔科夫定理知$\{\xi_k\}$满足大数定律。

中心极限定理不满足，这是因为

$$P\left(\sum_{k=1}^{n}\xi_k=0\right)\geqslant P(\xi_1=0,\ \cdots,\ \xi_n=0)$$

$$=\prod_{k=1}^{n}P(\xi_k=0)\geqslant\prod_{k=1}^{\infty}P(\xi_k=0)$$

$$=\prod_{k=1}^{\infty}\left(1-\frac{1}{2^{2k}}\right)>0,$$

这与 $$P\left(\frac{1}{B_n}\sum_{k=1}^{n}\xi_k<x\right)\longrightarrow\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{1}{2}t^2}dt$$

相矛盾。

（3）的例子　设$\{\xi_k\}$是相互独立的随机变量序列，且

$$P(\xi_k=k^\lambda)=\frac{1}{2},\quad P(\xi_k=-k^\lambda)=\frac{1}{2}$$

$$(k=1,\ 2,\ \cdots),$$

其中$\frac{1}{2}\leqslant\lambda\leqslant 1$。

$\{\xi_k\}$服从中心极限定理．事实上，因为

$$E\xi_k=0,\quad D\xi_k=k^{2\lambda}\quad(k=1,\ 2,\ \cdots),$$

$$B_n^2=\sum_{k=1}^{n}D\xi_k=\sum_{k=1}^{n}k^{2\lambda}.$$

由于
$$\lim_{n\to\infty}\frac{n^{2\lambda+1}}{B_n^2}=\lim_{n\to\infty}\frac{(n+1)^{2\lambda+1}-n^{2\lambda+1}}{B_{n+1}^2-B_n^2}$$

$$=\lim_{n\to\infty}\frac{(n+1)^{2\lambda+1}-n^{2\lambda+1}}{(n+1)^{2\lambda}}=2\lambda+1,$$

所以当$n\to\infty$时

$$B_n^2\approx\frac{n^{2\lambda+1}}{2\lambda+1}.$$

注意到 $|\xi_k|=k^\lambda$,

所以 $\max\limits_{1\leqslant k\leqslant n}|\xi_k|=\max\limits_{1\leqslant k\leqslant n}k^\lambda\leqslant n^\lambda$,

且 $\lim\limits_{n\to\infty}\frac{n^\lambda}{B_n}=\lim\limits_{n\to\infty}\frac{n^\lambda}{\left(\frac{n^{2\lambda+1}}{2\lambda+1}\right)^{\frac{1}{2}}}=0$.

根据本款有关的定理*知$\{\xi_k\}$满足中心极限定理．但是不满足大数定律，这是因为$\{\xi_k\}$相互独立，$E\xi_k=0$，且

$$E\left(\frac{\xi_k^2}{n^2+\xi_k^2}\right)=\frac{k^{2\lambda}}{n^2+k^{2\lambda}}.$$

注意到$\frac{1}{2}\leqslant\lambda\leqslant 1$，故有

$$\sum_{k=1}^{n}\frac{k^{2\lambda}}{n^2+k^{2\lambda}}\geqslant\sum_{k=1}^{n}\frac{k}{n^2+n^2}=\frac{1}{2n^2}\sum_{k=1}^{n}k$$

$$= \frac{n+1}{4n},$$

可见 $$\sum_{k=1}^{n} E\left[\frac{(\xi_k - E\xi_k)^2}{n^2 + (\xi_k - E\xi_k)^2}\right] = \sum_{k=1}^{n} E\left(\frac{\xi_k^2}{n^2 + \xi_k^2}\right)$$

$$= \sum_{k=1}^{n} \frac{k^{2\lambda}}{n^2 + k^{2\lambda}} \nrightarrow 0.$$

由格涅坚科定理知$\{\xi_k\}$不满足大数定律。

进一步的讨论

从上述三个例子看到，对于$\{\xi_n\}$不是同分布的情形，大数定律与中心极限定理之间的关系是不确定的。中心极限定理成立的随机变量序列可以不服从大数定律；大数定律成立的随机变量序列也可以不服从中心极限定理。甚至对有些随机变量序列来说，大数定律和中心极限定理两者可以都不满足。然而当$\{\xi_k\}$独立，同分布，有大于0的有穷方差时，大数定律和中心极限定理都成立。此时，大数定律断定，对任意$\varepsilon > 0$。

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n}(\xi_k - E\xi_k)\right| \leqslant \varepsilon\right) = 1.$$

然而括号中事件的概率有多大？此定律未能回答，但中心极限却给出一近似解答：

$$P\left(\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_k - E\xi_k)\right| \leqslant \varepsilon\right)$$

$$= P\left(\frac{1}{\sigma\sqrt{n}}\left|\sum_{k=1}^{n}(\xi_k - E\xi_k)\right| \leqslant \frac{\varepsilon\sqrt{n}}{\sigma}\right)$$

$$\approx \frac{1}{\sqrt{2\pi}}\int_{|x| < \frac{\varepsilon\sqrt{n}}{\sigma}} e^{-x^2/2}dx.$$

其中$\sigma^2 = D\xi_k$。因而在所假定的条件下，中心极限定理比大数定律更精确。

7—10 随机变量序列$\{\xi_n\}$独立同分布，但不服从中心极限定理的例子

有关概念

定理 设$\{\xi_n\}$为相互独立同分布的随机变量序列，且

$$E\xi_k = a, \quad D\xi_k = \sigma^2 < \infty$$

$$(k = 1, 2, \cdots),$$

则$\{\xi_n\}$服从中心极限定理。

例 设随机变量序列$\{\xi_n\}$相互独立且服从同一哥西分布，其密度函数为

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2},$$

考虑

$$\eta_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k.$$

由于ξ_k的特征函数为

$$\varphi_{\xi_k}(t) = e^{-|t|},$$

因此，η_n的特征函数为

$$\varphi_{\eta_n}(t) = e^{-n\frac{|t|}{n}} = e^{-|t|}.$$

这个结果表明，对任意n，η_n仍服从哥西分布，因此，当$n\to\infty$时，η_n的极限分布不是正态分布，故随机变量序列$\{\xi_n\}$的前n项平均数$\frac{1}{n}\sum_{k=1}^{n}\xi_k$的分布函数不能用正态分布来近似。因此$\{\xi_n\}$

不服从中心极限定理．究其原因，虽然$\{\xi_n\}$独立同分布，但是作为哥西分布，我们知道它的数学期望和方差都不存在，因此它不满足本款定理的条件．

(*) 7—11　$\mathbf{E}|\xi|^2=\infty$，Lindeberg-Ie′vy 中心极限定理仍成立的例子

有关命题

Lindeberg-Ie'vy 中心极限定理　设ξ_1，ξ_2，…，ξ_n，…是独立同分布的随机变量序列，

$$E\xi_k=\mu,\quad D\xi_k=\sigma^2,\quad (k=1,\ 2,\ \cdots),$$

则
$$S_n=\frac{1}{\sqrt{n}\,\sigma}\sum_{k=1}^{n}(\xi_k-\mu)\xrightarrow{L}Z.$$

其中Z是$N(0.1)$变量．

但$E|\xi|^2<\infty$仅是 Lindeberg-Ie'vy 中心极限定理成立的充分条件而不是必要条件，请看

反例

如果ξ_1，ξ_2，…，ξ_n，…是独立同分布的随机变量序列，$\xi_1\sim L(x)$使得

$$P\{|\xi|>x\}=\frac{L(x)}{x^{\zeta}},\quad \zeta\geqslant 2.$$

其中L是缓变函数$\left(\text{即}\dfrac{L(cx)}{L(x)}\xrightarrow[x\to\infty]{}1\text{对任意}c>0\right)$则$\xi_i$属于正态分布吸引场（见 V, K Rohatgi：Commun stat 2(1973)p525—533)，例如，若

$$\xi\sim f(x)=\begin{cases}2|x|^{-3}\log|x|, & \text{当}|x|\geqslant 1,\\ 0, & \text{当}|x|\leqslant 1.\end{cases}$$

则 $\frac{S_n}{\sqrt{2n\log n}} \xrightarrow{L} Z.$

其中 $S_n = \sum_{k=1}^{n} X_k$, Z是$N(0,1)$变量

(见〔13〕). 但$E|\xi|^2 = \infty$

(*) 7—12 $\{\xi_{nk} 1 \leqslant k \leqslant n,\ n \geqslant 1\}$满足$u$、$a$、$n$条件，但$\max\limits_{1\leqslant k\leqslant n} |\xi_{nk}| \overset{P}{\not\longrightarrow} 0$的例子

有关定义与定理

定义 设 $\xi_{11}, \cdots, \xi_{1k_1}$

$\xi_{21}, \cdots, \xi_{2k_2}$

…………

$\xi_{n1}, \cdots, \xi_{nk_n}$

是一串随机变量序列，若它们满足

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant k_n} P\{|\xi_{nk}| \geqslant \varepsilon\} = 0 \quad 对\forall \varepsilon > 0,$$

则称 $\{\xi_{nk}\}$ 满足u、a、n（Uniformly asymptotically negligible）条件。

定理1 设$\{\xi_{nk}\}$是一串行独立的随机变量序列，则条件

$$\sum_{k=1}^{k_n} P\{|\xi_{nk}| \geqslant \varepsilon\} \longrightarrow 0 \Longleftrightarrow \max_{1\leqslant k\leqslant k_n} |\xi_{nk}| \xrightarrow{P} 0$$

(见〔1〕P314注5.3.3)。

由于 $\max\limits_{1\leqslant k\leqslant k_n} P\{|\xi_{nk}| \geqslant \varepsilon\} \leqslant \sum\limits_{k=1}^{k_n} P\{|\xi_{nk}| \geqslant \varepsilon\}$.

由此不难由定理1得到

定理2 若

$$\max_{1\leqslant k\leqslant k_n}|\xi_{nk}|\xrightarrow{P}0,$$

则$\{\xi_{nk}\}$满足u、a、n条件.

但定理2的逆不真，请看

反例

设$\{\xi_{nk}\}1\leqslant k\leqslant n$，$n\geqslant 1$是一串行独立的随机变量序列，$\xi_{nk}$服从二点分布

$$P(\xi_{nk}=0)=1-\frac{1}{n},\ P(\xi_{nk}=1)=\frac{1}{n}(1\leqslant k\leqslant n),$$ 则对

$\forall\varepsilon$ $(0<\varepsilon<1)$ 有

$$\max_{1\leqslant k\leqslant n}P\{|\xi_{nk}|\geqslant\varepsilon\}=P\{\xi_{nk}=1\}=\frac{1}{n}\xrightarrow[(n\to\infty)]{}0,$$

即$\{\xi_{nk}\}$满足u、a、n条件，但对$0<\varepsilon<1$,

$$P\{\max_{1\leqslant k\leqslant n}|\xi_{nk}|\geqslant\varepsilon\}=1-P\{\max_{1\leqslant k\leqslant n}|\xi_{nk}|<\varepsilon\}$$

$$=1-\prod_{k=1}^{n}P\{|\xi_{nk}|<\varepsilon\}=1-\prod_{k=1}^{n}P\{\xi_{nk}=0\}$$

$$=1-\left(1-\frac{1}{n}\right)^n\longrightarrow 1-e^{-1}\neq 0,$$

即 $$\max_{1\leqslant k\leqslant n}|\xi_{nk}|\ \overset{P}{\not\longrightarrow}0.$$

(*) 7-13 $\sum_{n=1}^{\infty}\frac{D\xi_n}{n^2}=\infty$，$\{\xi_n\}$仍服从强大数定律的例子

有关定理

Kolmogorov 强大数定律 见例7—5，但

$$\sum_{n=1}^{\infty}\frac{D\xi_n}{n^2}<\infty$$

只是$\{\xi_n, n\geqslant 1\}$服从强大数定律之充分条件，而不是必要条件，请看

反例

设$\{\xi_n, n\geqslant 1\}$为独立的随机变量序列，满足

$$P\left\{\xi_n=\pm\sqrt{\frac{n}{\log(n+1)}}\right\}=\frac{1}{2},$$

则
$$E\xi_n=0,\quad D\xi_n=E\xi_n^2=\frac{n}{\log(n+1)},$$

$$E\xi_n^4=\frac{n^2}{\log^2(n+1)},$$

从而
$$\sum_{n=1}^{\infty}\frac{D\xi_n}{n^2}=\sum_{n=1}^{\infty}\frac{1}{n\log(n+1)}=\infty,$$

但
$$\sum_{n=1}^{\infty}\frac{E\xi_n^4}{n^3}=\sum_{n=1}^{\infty}\frac{1}{n\log^2(n+1)}<\infty,$$

于是由*H, D, Brunk—K, L Chung* 强大数定律（见〔19〕*P*333）知$\{\xi_n, n\geqslant 1\}$服从强大数定律。

(*) 7—14 $\sum_{n=1}^{\infty}(\xi_n-E\xi_n)$ *a*、*s*收敛，但 $\sum_{n=1}^{\infty}D\xi_n=\infty$的例子

有关命题

定理1 （Kolmogorov三级数定理）设$\{\xi_n\}$是独立随机变

量序列，则$\sum_{n=1}^{\infty}\xi_n a$、s收敛的必要和充分条件为对某个$C>0$

(i) $\sum_{n=1}^{\infty}P\{|\xi_n|\geqslant C\}<\infty$；

(ii) $\sum_{n=1}^{\infty}E\xi_n^c<\infty$；

(iii) $\sum_{n=1}^{\infty}D\xi_n^c<\infty$。

而且若(i) (ii) (iii)对某个$c>0$成立，则对任意的$c>0$也成立。

其中 $\xi_n^c=\begin{cases}\xi_n, & |\xi_n|<C,\\ 0, & |\xi_n|\geqslant C.\end{cases}$ （见〔1〕P88命题2.3.6）

定理2 若$\{\xi_n\}$是一串独立的随机变量序列，且$\sigma_n^2=D\xi_n<\infty$，$n=1, 2, \cdots$

则当$\sum_{n=1}^{\infty}\sigma_n^2<\infty$时，有$\sum_{n=1}^{\infty}(\xi_n-E\xi_n)a$、s收敛（见〔1〕P86命题2.3.3）。但定理2的逆不真，请看

反例

设$\{\xi_n, n\geqslant1\}$为独立的随机变量序列

$$P(\xi_n=\pm n)=\frac{1}{2n^2},\quad P(\xi_n=0)=1-\frac{1}{n^2},$$

则 $E\xi_n=0$，$\sigma_n^2=D\xi_n=1$，

$\therefore$ $\sum_{n=1}^{\infty}D\xi_n=\sum_{n=1}^{\infty}\sigma_n^2=\infty$。

但易验证$\{\xi_n\}$满足 Kalmogorov 三级数定理的充分性条件，故仍有

$$\sum_{n=1}^{\infty}(\xi_n - E\xi_n) = \sum_{n=1}^{\infty}\xi_n \quad a、s收敛。$$

注 若$\{\xi_n, n\geqslant 1\}$为独立的随机变量序列，且满足$P\{|\xi_n|\leqslant v$，对某个$v>0$，$\forall n\geqslant 1\}=1$，

则 $\sum_{n=1}^{\infty}(\xi_n - E\xi_n)a$、$s$收敛时，必有

$$\sum_{n=1}^{\infty}D\xi_n = \sum_{n=1}^{\infty}\sigma_n^2<\infty$$

（见命题2.3.4）。

第八章　与充分统计量和完全统计量有关的反例

统计量的充分性这个重要的基本概念，它是*R. A. Fisher*在其1922年的奠基性工作《On the Mathematical Foundations of statistics》中正式提出来的，但这个重要概念，Fisher早在1920年以前的几年里就开始形成了。当时他与天文学家Eddington争论这样一个问题：

设X_1，X_2，…，$X_n \sim N(n,\ \sigma^2)$，要估计反映测量精度的σ。当时较常用的估计有两个：一个是样本方差

$$S_n^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

的平方根S_n^*，一个是绝对平均偏差

$$d = \sqrt{\frac{\pi}{2}}\ \frac{1}{n}\sum_{i=1}^{n}|X_i - \overline{X}|$$

（常数$\sqrt{\frac{\pi}{2}}$选择的理由是使d为σ的无偏估计）。

Fisher赞成S_n^*而Eddington赞成d。Fisher在1920年的一篇文章中谈了他的理由，其中一点是“S_n^*包含了样本中关于σ的全部信息，而d则否。”正是在这一提法中包含了充分性这个概念。

至于完全统计量这个重要概念，则是Lehmann和Scheffe'在1950年提出来的。

本章对充分统计量与完全统计量的有关反例，作尽可能详细的论述。

8—1 并非一切统计量都是充分统计量的例子

有关定义和定理

定义 设$X=(X_1, X_2, \cdots, X_n)$是从具有分布族为$\{F(x; \theta), \theta\in\Theta\}$的母体中抽取的一个子样，$T(X)$是一统计量（可以是向量），如果给定$T(X)=t$，$X$的条件分布（离散型变量为条件概率，连续型变量为条件密度）与参数θ无关，则称统计量$T(X)$是分布族$\{F(x; \theta), \theta\in\Theta\}$的充分统计量，或称$T(X)$是$\theta$的充分统计量.

直接用上述定义去验证一个统计量T的充分性，往往要经过复杂的计算. 幸好，我们有下面的一般性判定法，它在应用上很方便.

定理 (Fisher–Neyman 因子分解定理)

(1) 连续型情形 设母体X具有分布密度族$\{f(x; \theta), \theta\in\Theta\}$，$X=(X_1, X_2, \cdots, X_n)$是一子样，$T(X)$是一个统计量. 则$T(X)$为$\theta$的充分统计量的充要条件是：子样的联合分布密度函数可分解为

$$L(x; \theta)=\prod_{i=1}^{n} f(x_i; \theta)=h(x)g(T(x), \theta)$$

其中h是x的非负可测函数且与θ无关；$g(T(x), \theta)$仅通过T依赖于$x=(x_1, \cdots, x_n)$且对固定的θ，$g(t, \theta)$为t的非负可测函数.

(2) 离散型情形 若母体X具体概率分布族$\{p(x;\theta); \theta\in\Theta\}$，$X=(X_1, \cdots, X_n)$是一个子样，$T(X)$是一统计量，则$T(X)$为$\theta$的充分统计量的充要条件是，子样的联合概率分布可表示为

$$P(x;\theta)=h(x)g(T(x),\theta),$$

其中h是x的非负可测函数且与θ无关；$g(T(x),\theta)$ 仅通过T依赖于$x=(x_1,\cdots,x_n)$，且对固定的θ, $g(t,\theta)$ 是t的非负可测函数。

如果θ是参数向量，T是随机向量且定理的条件成立，则称T关于θ是联合充分的。

在一般教科书中，给定的统计量T都是充分统计量，但是，并非所有的统计量都是充分统计量。

注意到，Fisher-Neyman 因子分解定理一般用来判定一个给定的统计量T 是充分统计量是比较有效的，而用它来判定一个给定的统计量 T不是充分统计量往往不太方便，不如直接按充分统计量的定义来判别容易。我们要判断T不是充分统计量，按定义即要验证在固定 $T=t$下，$X=(X_1,\cdots,X_n)$的条件分布不依赖于要估计的参数θ的条件不满足。下面我们给出一个不是充分统计量的例子。

例1 设X_1,X_2是从总体$\boldsymbol{X}\sim p(\lambda)$(poisson 分布)中抽取的两个iid (independent identically distributions 独立同分布)样本，考虑统计量

$$T=X_1+2X_2$$

由于
$$p=(X_1=0,\ X_2=1|X_1+2X_2=2)$$
$$=\frac{P(X_1=0,X_2=1,X_1+2X_2=2)}{P(X_1+2X_2=2)}$$
$$=\frac{P(X_1=0,X_2=1)}{P(X_1+2X_2=2)}$$
$$=\frac{\lambda e^{-2\lambda}}{P(X_1=0,X_2=1)+P(X_1=2,X_2=0)}$$
$$=\frac{\lambda e^{-2\lambda}}{\lambda e^{-2\lambda}+\left(\frac{\lambda^2}{2}\right)e^{-2\lambda}}=\frac{1}{1+\left(\frac{\lambda}{2}\right)}\qquad \text{与}\lambda\text{有关}$$

故统计量$T=X_1+2X_2$。
不是参数λ的充分统计量。

进一步的讨论

注 但统计量

$$T=X_1+X_2$$

是参数λ的充分统计量。事实上

$$P(X_1=x_1,X_2=x_2|X_1+X_2=t)$$

$$=\begin{cases}\dfrac{P(X_1=x_1,X_2=t-x_1)}{P(X_1+X_2=t)} & \text{如果} t=x_1+x_2,\ x_i=0,1,2,\cdots,\ i=1,2, \\ 0, & \text{其他.}\end{cases}$$

所以，对于$x_i=0,1,2,\cdots$，$i=1,2$，$x_1+x_2=t$，我们有

$$P(X_1=x_1,X_2=x_2|X_1+X_2=t)$$

$$=\frac{\dfrac{\lambda^{x_1}}{x_1!}e^{-\lambda}\dfrac{\lambda^{t-x_1}}{(t-x_1)!}e^{-\lambda}}{\dfrac{(2\lambda)^t}{t!}e^{-2\lambda}}$$

$$=\frac{t!}{x_1!(t-x_1)!}\frac{1}{2^t}=\binom{t}{x_1}\left(\frac{1}{2}\right)^t \text{与λ无关，}$$

(*) 8—2 极小充分统计量不是完全充分统计量的例子

有关定义和命题

定义1 设$t(x)$是$(*,\ \beta_*)$上的一个充分统计量，取值于$(J,\ \beta_J)$，β_*上的分布族为$\{p_\theta,\ \theta\in\Theta\}$，若对任何定义于$*$，取值于某可测空间$(s,\ \beta_s)$的充分统计量$s(x)$，必存在由$(s,\ \beta_s)$到$(J,\ \beta_J)$的可测变换$t=q(s)$，以及$A\in B_*$满足条件

$P_\theta(A)=0$对任何$\theta\in\Theta$，使

$t(x)=q(s(x))$对任何$x\in A$

则称t为一极小充分统计量.

定义2 （完全分布族与完全统计量）*设$\{\rho_\theta,\theta\in\Theta\}$为样本空间（$*,\beta_*$）上的一族概率分布，若对任何定义于$*$上的$\beta_*$可测函数$f(x)$，满足条件

$E_\theta(f(X))=0$，对一切$\theta\in\Theta$，

即　　$\int_* f(x)d\rho_\theta(x)=0$，对一切$\theta\in\Theta$.

必有　　$f(X)\doteq 0$　　$(a,\ s\rho_\theta)$对任何$\theta\in\Theta$，

即　　$\rho_\theta(f(X)=0)=1$　对任何$\theta\in\Theta$.

则称分布族$\{\rho_\theta,\theta\in\Theta\}$是完全的.

若$\{\rho_\theta,\ \theta\in\Theta\}$为$(*,\beta_*)$上的一个概率分布族(不必为完全的)，$t(x)$为定义于$*$上，取值于$(J,\ \beta_J)$的统计量，$\{\rho_\theta^T,\theta\in\Theta\}$是$\{\rho_\theta,\ \theta\in\Theta\}$的导出分布族，若$\{\rho_\theta^T,\ \theta\in\Theta\}$为完全分布族，则称$t(x)$为一完全统计量. 因此，要验证统计量$t$的完全性，需要求出$t$的分布族$\{\rho_\theta^T,\ \theta\in\Theta\}$，然后证明，若$\int_J g(t)d\rho_\theta^T(t)=0$，对一切$\theta\in\Theta$

必有　　$\rho_\theta^T\{t,g(t)\neq 0\}=0$，对任何$\theta\in\Theta$

或　　$\rho_\theta^T\{t,g(t)=0\}=1$，对任何$\theta\in\Theta$.

我们有如下的正面命题：

定理　完全充分统计量必为极小充分统计量(见〔7〕p80).

但其逆不真，即极小充分统计量未必是完全充分统计量，反例如下：

例1　设样本空间为$*=(-\infty,\ \infty)$，参数空间为

$\Theta=\{\theta:\ \theta>1\}$.

---

* 或称完备分布族与完备统计量.

分布族为$\mathscr{F}=\{p_\theta,\ \theta>1\}$，

$$\left.\begin{aligned}&p_\theta(X=\theta)=p_\theta(X=-\theta)=\frac{1}{2\theta^2},\\&p_\theta(X=0)=1-\frac{1}{\theta^2}.\end{aligned}\right\}\theta>1$$

此分布关于0点对称。

不难证明$t(x)=|x|$的极小性。事实上，设$g(x)$为θ的任一充分统计量，只须证明$|x_1|=|x_2|$，可以推出$g(x_1)=g(x_2)$，则证明了t的极小性。用反证法，设$|a|\neq|b|$，但$g(a)=g(b)$（公共值记为c），又为确定起见，先设$a\neq0$，$b\neq0$。

考虑集合$A=\{x:\ g(x)=c\}$，则A中至少包含a、b两点。

当分布为$P_{|a|}$时，有

$$P_{|a|}(x=a|g(x)=c)=\begin{cases}1, & \text{当}A\text{不包含}0,\ -a,\\ \dfrac{1}{2a^2}, & \text{当}A\text{包含}0,\ -a,\\ \dfrac{1}{2}, & \text{当}A\text{包含}-a,\text{不包含}0,\\ \dfrac{1}{2a^2-1}, & \text{当}A\text{包含}0,\text{不包含}-a\end{cases}$$

（以上这些全不为0）

而当分布为$P_{|b|}$时，有

$$P_{|b|}(X=a|g(x)=c)=0.$$

因此给定$g(x)$时，X的条件分布与参数θ之值有关（例如θ取$|a|$和θ取$|b|$时，不同，如上述）故$g(x)$不是充分的，矛盾。

若a、b中有一个为0，证明类似。这样就证明了$t(x)=|x|$的极小性，即$t(x)=|x|$为极小充分统计量。但t并非完全。因为若取

$$h(t)=t^2,$$

则　　$E_{\theta}(h(t))=E_{\theta}(x^2)=D_{\theta}(x)=1$，

故若取　$H(t)=t^2-1$，

则　　$E_{\theta}(H(t))=0$　对一切$\theta>1$.

但$H(t)$显然不概率1为0，因此t并非完全，即t不是完全充分统计量.

8—3　次序统计量不是完全统计量的例子

有关定义和定理

定义1　设$X=(X_1,\cdots,X_n)$是从母体$\mathbf{X}$中抽取的一个子样，记$(x_1,\cdots,x_n)$是子样的一个观察值．将观察值的各分量按大小递增次序排列，得到

$$x_1^*\leqslant x_2^*\leqslant\cdots\leqslant x_n^*$$

当$(X_1,\cdots,X_n)$取值为$(x_1,\cdots,x_n)$时，我们定义$X_{(k)}$取值x_k^*，称由此得到的$(X_{(1)},\cdots,X_{(n)})$为$(X_1,\cdots,X_n)$的一组次序统计量．显然有$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$.

首先，我们有如下的事实：

定理　设分布族满足以下两个条件：

(a) 若$F_1\in\mathscr{F}_1$，$F_2\in\mathscr{F}$，则对任何$p_1>0$，$p_2>0$，$p_1+p_2=1$，有$p_1F_1+p_2F_2\in\mathscr{F}$

(b)若$F\in\mathscr{F}$，$s=[a,b)$　$-\infty<a<b<\infty$，而

$$F(s)\neq0(\text{即}F(s)=F(b)-F(a)>0)$$

则　　$F_s\in\mathscr{F}$

其中　$$F_s(x)=\begin{cases}\dfrac{F(x)-F(a)}{F(s)}, & \text{当}a<x<b,\\ 0, & \text{当}x\leqslant a,\\ 1, & \text{当}x\geqslant b.\end{cases}$$

则次序统计量（$X_{(1)}$，$X_{(2)}$，…，$X_{(n)}$）是完全的（见[7]p83）。

但如果将定理中的(a)、(b)去掉一个，则定理的结论不再成立，我们有如下的反例。

例1 设X_1，X_2，…，X_n为iid样本，$x_1 \sim N(0, \sigma^2)$，显然$N(0, \sigma^2)$不满足定理之条件(a)

令 $$t(X_{(1)}, \cdots, X_{(n)}) = \sum_{i=1}^{n} X_{(i)} = \sum_{i=1}^{n} X_i,$$

则 $$E_{\sigma^2}(t(X_{(1)}, \cdots, X_{(n)})) = 0,$$

但 $$t(X_{(1)}, \cdots, X_{(n)}) = 0 \quad a.e \ \underbrace{F \times F \times \cdots \times F}_{n}$$ 不成立，其中F为x_1之分布函数。

这说明$t(X_{(1)}, \cdots, X_{(n)})$不是完全统计量。

例2 分布族仅含有一个分布$U(0,1)$，（(0,1)上的均匀分布），记

$$U(0,1) \triangleq F,$$

易知F不满足定理之条件(b)。

对此分布族而言，次序统计量不是完全的。

例如 $$\prod_{i=1}^{n}\left(X_{(i)} - \frac{1}{2}\right) = \prod_{i=1}^{n}\left(X_i - \frac{1}{2}\right),$$

有 $$E\left(\prod_{i=1}^{n}\left(X_{(i)} - \frac{1}{2}\right)\right) = \int_0^1 \cdots \int_0^1 \prod_{i=1}^{n}\left(x_i - \frac{1}{2}\right) dx_1 \cdots dx_n = 0,$$

但 $$\prod_{i=1}^{n}\left(X_{(i)} - \frac{1}{2}\right) = 0 \qquad a.e \ \underbrace{F \times F \times \cdots \times F}_{n}$$ 不成立。

进一步的讨论

注 定理可推广到多组样本的情况，我们有

定理 设X_{i_1}，…，X_{in_i}为抽自分布族为F_i的母体的iid样本，$i=1$，…，c又X_{11}，…，X_{1n_1}，…，X_{c1}，…，X_{cn_c}全体独立，假定F_i属于分布族$\mathscr{F}_i$，$i=1$，…，c，以$X_{i(1)}\leqslant X_{i(2)}\leqslant\cdots\leqslant X_{i(n_i)}$记$X_{i1}$，…，$X_{in_i}$的次序统计量，$i=1$，…，$c$，而定义

$$T(X_{11},\ \cdots,\ X_{1n_1},\ \cdots,\ X_{c1},\ \cdots X_{cn_c})$$
$$=(X_{1(1)},\ \cdots,\ X_{1(n_1)},\ \cdots,\ X_{C(1)},\ \cdots X_{C(n_c)}).$$

如果每个$\mathscr{F}_i$，$i=1$，…，c都满足上一定理的条件（a）、（b），则T是完全统计量。

完全类似，此定理的逆不真。

8—4 有界完全的分布族（或统计量）不是完全的分布族（或统计量）的例子

有关定义和定理

定义 设随机变量X的样本空间为（$*_*$ β_*），分布族为$\{\rho_\theta,\ \theta\in\Theta\}$，$t(x)$是定义于$*$上，取值于$(J,\beta_J)$的统计量，其分布族为$\{\rho_\theta^T,\ \theta\in\Theta\}$，若对任何满足条件

$\int_* f(x)d\rho_\theta(x)=0$，对一切$\theta\in\Theta$

的有界、β_*可测函数$f(x)$，必有

$f(x)=0$，$(a_\text{、}e\rho_\theta)$对任何$\theta\in\Theta$，

则称分布族$\{\rho_\theta,\ \theta\in\Theta\}$为有界完全的。若$\{p_\theta^T,\ \theta\in\Theta\}$为有界完全的，则称$t(x)$为有界完全统计量。由完全统计量的定义及有界完全统计量的定义，显然有如下的

定理 完全的分布族或统计量，必为有界完全的。

但其逆不成立，我们有如下的

例1 设 $* = \{0,1,2,\cdots\}$，分布族为：

$$p_\theta(x=0)=\theta,$$

$$p_\theta(x=n)=(1-\theta)^2\theta^{n-1},\quad n=1,2,\cdots,\quad 0<\theta<1.$$

设 $E_\theta(f(X))=0$，对一切θ，$0<\theta<1$，

则
$$f(0)p_\theta(x=0)+\sum_{n=1}^{\infty}f(n)p_\theta(x=n)=0,0<\theta<1.$$

也即
$$f(0)\theta+(1-\theta)^2\sum_{n=1}^{\infty}f(n)\theta^{n-1}=0,0<\theta<1.$$

于是
$$\sum_{n=1}^{\infty}f(n)\theta^{n-1}=-f(0)\frac{\theta}{(1-\theta)^2}=-f(0)\sum_{n=1}^{\infty}n\theta^n$$
$$0<\theta<1.$$

上式两边为θ的幂级数，在$|\theta|<1$内收敛，故其对应项系数必相同，即

$$f(1)=0,$$

$$f(n)=-(n-1)f(0),\quad n=2,3,\cdots$$

若要求f有界，则由此知必须有$f(0)=0$。因而

$$f(1)=f(2)=f(3)=\cdots=0,$$

这证明了 $\{p_\theta,\ \theta\in(0,1)=\Theta\}$ 为有界完全的分布族。然而，$\{p_\theta,\ \theta\in(0,1)=\Theta\}$不是完全的，因为，若取

$$f(0)=-1,$$

$$f(1)=0,$$

$$f(n)=n-1,\quad n=2,3,\cdots$$

则
$$E_\theta(f(X))=f(0)p_\theta(X=0)+f(1)p_\theta(X=1)+\sum_{n=2}^{\infty}f(n)p_\theta(X=n)$$

$$=-\theta+(1-\theta)^2\sum_{n=2}^{\infty}(n-1)\theta^{n-1}\cdot$$

$$= -\theta + (1-\theta)^2 \sum_{n=1}^{\infty} n\theta^n$$

$$= -\theta + (1-\theta)^2 \frac{\theta}{(1-\theta)^2}.$$

$$= 0, \quad \text{对一切}\theta \in (0,1) = \Theta.$$

但 $p_\theta(f(X) = 0)$ 并不为1.

因为 $p_\theta(f(X) = 0) = p_\theta(X = 1) = (1-\theta)^2 < 1 \quad (0 < \theta < 1)$.

进一步的讨论

注 不完全的分布族也可以存在完全的统计量.请见下例:

例2 设 $X = (X_1, \cdots, X_n)$ 为取自贝努里 $b(1;\theta)$ 分布的 iid 样本，取统计量

$$t(X) = \sum_{i=1}^{n} X_i.$$

则 t 的分布为二项分布 $b(n;\ \theta)$ $0 \leqslant \theta \leqslant 1$.

$$J = (0,1,2,\cdots,n).$$

$$p_\theta^T(t = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, \quad k = 0,1,2,\cdots,n.$$

这时条件

$$\int_J g(t) P_\theta^T(t) = 0, \qquad 0 \leqslant \theta \leqslant 1$$

成为 $\displaystyle\sum_{k=0}^{n} g(k)\binom{n}{k}\theta^k(1-\theta)^{n-k} = 0, \quad 0 \leqslant \theta \leqslant 1.$

更有 $\displaystyle\sum_{k=0}^{n} g(k)\binom{n}{k}\theta^k(1-\theta)^{n-k} = 0, \quad 0 < \theta < 1.$

因而，若记 $y = \dfrac{\theta}{1-\theta}$，则当 $0 < \theta < 1$ 时，有 $0 < y < \infty$，

而 $$\sum_{k=0}^{n} g(k)\binom{n}{k} y^k = 0,\quad 0<y<\infty.$$

这是y的n次多项式，它对任意的$0<y<\infty$都等于0，而这必须每项系数

$$\binom{n}{k} g(k) = 0,\quad k = 0,1,2,\cdots,n.$$

即 $$g(k) = 0,\quad k = 0,1,2,\cdots,n.$$

这证明了t为完全统计量.

但贝努里分布族为不完全分布族. 这是因为贝努里变量之样本空间

$$\mathscr{X} = \{(x_1,\cdots,x_n),\ x_i\text{取}0\text{或}1\}.$$

$\mathscr{X}$中的样本点为2^n个；β_*为$\mathscr{X}$的一切子集构成的σ-域. 分布族为

$$\begin{aligned}P_\theta(X = x) &= P_\theta(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)\\ &= \theta^k(1-\theta)^{n-k},\end{aligned}$$

其中k为$x_1,\cdots,x_n$中取1的个数.

但如取$f(X) = X_1 - X_2$，这时

$$E_\theta(f(X)) = EX_1 - EX_2 = \theta - \theta = 0,$$

但$f(X) = 0$的概率不为1，这是因为

$$\begin{aligned}P_\theta(f(X) = 0) &= P_\theta(X_1 - X_2 = 0)\\ &= P_\theta(X_1 = 0,\ X_2 = 0) + P_\theta(X_1 = 1,\ X_2 = 1)\\ &= (1-\theta)^2 + \theta^2 < 1 - \theta + \theta = 1 \quad (0<\theta<1).\end{aligned}$$

所以，此分布族为不完全分布族.

注 对正态，Poisson…等分布族也有类似的情况.

(*)8—5 可测函数$f(X)$与有界充分完全统计量$t(X)$独立，但$f(X)$的分布与θ有关的例子

有关定理

首先，我们有如下的结论：

定理 设随机变量X的样本空间和分布族为（$\mathscr{X}$，$\beta_{\mathscr{X}}$）及$\{P_\theta,\ \theta\in\Theta\}$，而$t(x)$为一有界完全统计量，取值于$(J,\ \beta_J)$，且 $t(x)$ 为充分统计量．则对任何定义于$\mathscr{X}$的有限$\beta_{\mathscr{X}}$—可测函数 $f(x)$，当 $f(X)$ 的分布与θ无关时，对一切$\theta\in\Theta$，$f(X)$ 与 $t(X)$ 独立（见〔7〕P88）．然而，值得注意的是，本定理之逆不真，请看：

例1 设$\mathscr{X}=\Theta=J=R_1$，$\beta_{\mathscr{X}}=\beta_J=\beta^1$．

而X之分布族为

$$P_\theta(\{\theta\})=1,\ \theta\in R_1.$$

则可证$t(x)=x$为一完全充分统计量．

令$f(x)=x$，则对任何$\theta\in R_1$，

$$P_\theta(t(x)=\theta)=P_\theta(f(x)=\theta)=1.$$

即对任何$\theta\in R_1$，$f(X)$和$t(X)$在任何$P_\theta(\theta\in R_1)$之下都是退化分布，所以，对任何$\theta\in R_1$，有

$$P_\theta(t(x)=\theta,\ f(x)=\theta)=1=P_\theta(t(x)=\theta)P_\theta(f(x)=\theta).$$

即对任何$\theta\in R_1$，$f(X)$与$t(X)$独立，但由X之分布族为

$$P_\theta(\{\theta\})=1,\ -\infty<\theta<\infty.$$

知$f(X)$之分布并非与θ无关．

8—6 充分统计量的函数不是充分统计量的例子

有关定理

众所周知，我们有如下的

定理 设 $T(X)$ 是θ的充分统计量，$v=\psi(t)$是单值可逆函数，则$v=\psi(T)$ 也是θ的充分统计量。

但这并不意味着任何时候，任一充分统计量的函数它本身也是充分统计量。请看

例1 设$\boldsymbol{X}\sim N(\mu,\ \sigma^2)$，$\sigma^2$已知。

$X=(X_1,\cdots,X_n)$ 是抽自$\boldsymbol{X}$的iid 样本，则依因子分解定理可知$\overline{X}$是μ的充分统计量。但 $\overline{X}^2$不是μ的充分统计量，事实上

$$f(x_1,\cdots,x_n\,|\,\overline{X}^2=t)=\frac{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}}{f(t)}$$

$$=\frac{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}}{\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\left[e^{-\frac{n(\overline{X}+\mu)^2}{2\sigma^2}}+e^{-\frac{n(\overline{X}-\mu)^2}{2\sigma^2}}\right]}\text{ 与}\mu\text{有关}$$

故$\overline{X}^2$不是μ的充分统计量。

$$\because\quad \overline{X}\sim\varphi(t)=N\left(\mu,\ \frac{\sigma^2}{n}\right)=\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{n(t-\mu)^2}{2\sigma^2}},$$

$$\therefore\quad \overline{X}^2\sim f(t)=\varphi(-\sqrt{t})\left|-\frac{1}{2\sqrt{t}}\right|+\varphi(\sqrt{t})\left|\frac{1}{2\sqrt{t}}\right|$$

$$=\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{n(-\sqrt{t}-\mu)^2}{2\sigma^2}}\frac{1}{2\sqrt{t}}+\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{n(\sqrt{t}-\mu)^2}{2\sigma^2}}.$$

8—7 充分完全统计量的函数不是充分完全统计量的例子

有关定理

同样，我们有如下

定理 设$T(X)$是θ的充分完全统计量，$S(T)$是T的单值可逆函数，则$S(T)$也是θ的充分完全统计量。

但并不是任何充分完全统计量的函数都是充分完全统计量，例子如下

例1 设$X_1,X_2,\cdots,X_n$为iid样本，$x_1\sim N(\mu,\ \sigma^2)$，$\sigma^2$已知，则$\overline{X}$是$\mu$的充分完全统计量。

虽然$\overline{X}^2$是μ的完全统计量（不难证明完全统计量的任一函数都是完全统计量），但由于$\overline{X}^2$不是μ的充分统计量，因而$\overline{X}^2$不是μ的充分完全统计量。

8—8 指数族分布表达式中的$T(X)$不是充分完全统计量的例子

有关定义和定理

定义 如果分布密度族$\{f(x;\ \theta),\ \theta\in\Theta\}$，其中$f(x;\ \theta)$有形式：

$$f(x;\ \theta)=a(\theta)\exp\left\{\sum_{j=1}^{k}Q_j(\theta)T_j(x)\right\}h(x),$$

其中$h(x)$是非负可测函数。则称它为指数族分布。

对指数族分布，我们有如下的

定理 设$\{f(x;\ \theta),\ \theta\in\Theta\}$是$k$个参数的指数族，分布密度为

$$f(x;\ \theta)=a(\theta)\exp\left\{\sum_{j=1}^{k}Q_j(\theta)T_j(x)\right\}h(x),$$

其中参数$\theta=(\theta_1,\cdots,\theta_k)\in\Theta$，如果$\Theta$包含一个$k$维矩形，且$Q=(Q_1,\ Q_2,\ \cdots,\ Q_k)$的值域包含有一个$k$维开集，则$T(X)=(T_1(X),\cdots,T_k(X))$是$\theta$的充分完全统计量（见〔6〕第二册之第二分册p170）。

但是，我们可以给出一个指数族分布，其中$T(X)$并不是参数θ的充分完全统计量，请看

例1 设母体$\boldsymbol{X}\sim N(\sigma,\ \sigma^2)$，其中$\sigma>0$是未知参数，$X=(X_1,\cdots,X_n)$是抽自其中的$iid$样本，则

$$T(X)=\left(\sum_{i=1}^{n}X_i,\ \ \sum_{i=1}^{n}X_i^2\right)$$

是σ的充分统计量，但不是σ的完全统计量，事实上，因为$\boldsymbol{X}\sim N(\sigma,\ \sigma^2)$，所以，子样$X=(X_1,\cdots,X_n)$的联合分布为

$$L(x;\ \theta)=\prod_{i=1}^{k}f(x_i;\ \theta)$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\sigma)^2\right\}$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2\right.$$

$$\left.+\frac{1}{\sigma}\sum_{i=1}^{n}X_i-\frac{n}{2}\right\}.$$

令$h(x)=1$,

$$g(T(x),\ \theta)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2\right.$$

$$\left.+\frac{1}{\sigma}\sum_{i=1}^{n}X_i-\frac{n}{2}\right\},$$

则 $$L(x;\ \theta)=g(T(x),\theta)h(x),$$

于是，根据Fisher-Neyman因子分解定理得

$$T(X)=\left(\sum_{i=1}^{n}X_i,\ \sum_{i=1}^{n}X_i^2\right).$$

是σ的充分统计量，记

$$T=\left(\sum_{i=1}^{n}X_i,\ \sum_{i=1}^{n}X_i^2\right)=(U,\ V).$$

则 $$EU=\sum_{i=1}^{n}EX_i=n\sigma,$$

$$DU=\sum_{i=1}^{n}DX_i=n\sigma^2,$$

∴ $$EU^2=DU+(EU)^2=n(n+1)\sigma^2.$$

又 $$EV=\sum_{i=1}^{n}EX_i^2=\sum_{i=1}^{n}[DX_i+(EX_i)^2]=2n\sigma^2,$$

令 $$g(T)=2U^2-(n+1)V,$$

则 $$Eg(T)=2n(n+1)\sigma^2-(n+1)2n\sigma^2=0.$$

但$n\geqslant 2$时，$g(T)\neq 0$。故

$$T=\left(\sum_{i=1}^{n}X_i,\ \sum_{i=1}^{n}X_i^2\right).$$

不是σ的完全统计量.

为什么会出现这一情况呢，原因是上述定理中

$$Q=(Q_1, Q_2)=\left(-\frac{1}{2\sigma^2}, \frac{1}{\sigma}\right).$$

之值域应包含一个二维开集的条件得不到满足。

进一步的讨论

注 设$X_1,\cdots,X_n$是抽自$N(a\sigma, \sigma^2)$分布族的iid样本，其中a是已知的实数，$\sigma>0$是未知参数，则同样可知

$$T(X)=\left(\sum_{i=1}^{n}X_i, \sum_{i=1}^{n}X_i^2\right)$$

是σ的充分统计量，但不是σ的完全统计量。

注 设$X\sim N(\mu, \sigma_1^2)$，$Y\sim N(\mu, \sigma_2^2)$
其中$(\mu, \sigma_1^2, \sigma_2^2)$是未知参数，又设

$$(X_1, X_2, \cdots, X_m), (Y_1, Y_2, \cdots, Y_n)$$

是分别从X、Y母体中抽取的iid样本，则可证明

$$T(X, Y)=\left(\sum_{i=1}^{m}X_i, \sum_{j=1}^{n}Y_j, \sum_{i=1}^{m}X_i^2, \sum_{j=1}^{n}Y_j^2\right).$$

是$(\mu, \sigma_1^2, \sigma_2^2)$的充分统计量，但不是完全统计量。原因是上述定理中

$$Q=(Q_1,Q_2,Q_3,Q_4)=\left(\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, -\frac{1}{2\sigma_1^2}, -\frac{1}{2\sigma_2^2}\right)$$

的值域应包含一个4维开集的条件得不到满足。

(*)8—9 在适当的条件下，当$T_i(X_i)$是θ_i($i=1, 2$)的充分统计量时，($T_1(X_1)$, $T_2(X_2)$)是(θ_1, θ_2)的联合充分统计量的例子

在 Fisher–Neymam 因子分解定理中，如果θ和T的维数相同，我们并不能由T关于θ的联合充分性而推出T的第j个分量关于θ的第j个分量是充分的. 请看

例1 设 $X_1, X_2, \cdots, X_n$为 iid 样本，$X_1 \sim N(\mu, \sigma^2)$，其中$\mu$，$\sigma^2$为未知参数. 则

$$X=(X_1, X_2, \cdots, X_n) \sim f(x; \mu, \sigma^2)$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}X_i^2\right.\right.$$

$$\left.\left.+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}X_i-\frac{n\mu^2}{2\sigma^2}\right)\right\}.$$

由因子分解定理，知统计量

$$T(X_1, X_2, \cdots, X_n)=\left(\sum_{i=1}^{n}X_i, \quad \sum_{i=1}^{n}X_i^2\right)$$

是(μ, σ^2)的联合充分统计量，或等价地

$$T^*(X_1, X_2, \cdots, X_n)=\left(\overline{X}, \quad \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right)$$

是(μ, σ^2)的联合充分统计量.

但是，当σ^2未知时，T^*之分量$\overline{X}$不是(μ, σ^2)的分量μ的充

分统计量。当μ未知时，T^*之分量$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$不是$(\mu,\sigma^2)$的分量$\sigma^2$的充分统计量。然而，当$\sigma^2$已知时，$\overline{X}$是$\mu$的充分统计量，当$\mu=\mu_0$已知时，$\sum_{i=1}^{n}(X_i-\mu_0)^2$ 是σ^2的充分统计量。尽管，由T关于θ的联合充分性，推不出T的第j个分量关于θ的第j个分量的充分性，但在适当的条件下，其逆为真，即我们有如下的

定理 如果X_i服从分布族$(\mathscr{X}_i,\ \mathscr{L}_i,\ p_{\theta_i}^{X_i},\ \theta_i\in\Theta_i)$ $(i=1,2)$，若$t_i(x_i)$是$\theta_i\in\Theta_i$的充分统计量$(i=1,2)$，则$(t_1(x_1),t_2(x_2))$是$(\theta_1,\theta_2)\in\Theta_1\times\Theta_2$的充分统计量（参看 D A. S. Fraser: Nonparametric Methods in Statistics John Wiley. New York 1956 p21)。

第九章　与点估计有关的反例

点估计是统计推断的基本问题之一，它是数理统计中一个内容比较丰富，反例也比较多的重要分支.我们分成无偏估计，Rao—cramer 不等式，可容许估计，大样本理论的基本概念，极大似然估计等五个方面提供它们的有关反例，并给出证明.

9—1　参数不存在无偏估计的例子

有关概念

定义　设母体 X 具有分布族 $\{F(x;\ \theta),\ \theta\in\Theta\}$，设 $X=(X_1,\cdots,X_n)$是从该母体中抽取的一个子样，$T(X)$是θ的一个估计，如果它满足下面的关系式

$$E_\theta(T)=\theta,\ \text{对一切}\theta\in\Theta,$$

则称$T(X)$是参数θ的一个无偏估计.

我们知道，若$X\sim N(\theta,\ \sigma^2)$，$X_1,X_2,\cdots,X_n$为其$iid$样本，则

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

为参数θ之无偏估计，

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

为参数 σ^2 之无偏估计．同样，若 $X\sim b(n;\ p)$（二项分布），$0<p<1$，$X_1,\cdots,X_n$为其iid样本，则$\frac{1}{n^2}\sum_{i=1}^{n}X_i$是参数 p 的无偏估计．但是，并非在任何情况下，参数的无偏估计都存在，请看

例1 设$X\sim N(\theta,\ 1)$，欲从一组iid样本$X_1,X_2,\cdots,X_n$出发估计参数$g(\theta)=|\theta|$．

设$\hat{g}(X)$为$g(\theta)=|\theta|$之无偏估计，则对$-\infty<\theta<\infty$，有

$$E_\theta(\hat{g}(X))=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}g(x)e^{-\frac{(x-\theta)^2}{2}}dx=|\theta|.$$

但由指数族的性质（见〔7〕p21之定理1、2、1）知上式中间一项，在$-\infty<\theta<\infty$，对θ有各级连续导数，而上式右边一项$|\theta|$在$\theta=0$处不可导，因此是不可能的，即对参数$|\theta|$而言，其无偏估计不存在．

例2 设$X\sim b(n;\ p)$，$0<p<1$．欲从一组iid样本出发，估计$g(p)=\frac{1}{p}$．

若$\hat{g}(X)$为$\frac{1}{p}$的无偏估计，记$\hat{g}(i)=a_i$，$i=0,1,\cdots,n$，则将有

$$E\hat{g}(X)=\sum_{i=0}^{n}\hat{g}(i)\binom{n}{i}p^i(1-p)^{n-i}$$

$$=\frac{1}{p},\quad 0<p<1.$$

这显然是不可能的，因为上式左边是 p 的次数不超过 n 的多项式，而右边在$p=0$处没有意义．

进一步的讨论

注 显然在例2中，只要$g(p)$不是p的次数$\leqslant n$的多项式，则$g(\theta)$的无偏估计都不可能存在。

9—2 无偏估计不是一致最小方差无偏估计的例子

有关定义和定理

定义 记$U \triangleq \{T: E_\theta T=\theta, E_\theta T^2<\infty$，对一切$\theta\in\Theta$成立$\}$即$U$是具有有限方差的$\theta$的全体无偏估计所组成的集合。如果对$T_0\in U$，成立下式

$$D_\theta(T_0)\leqslant D_\theta(T) \quad \text{对一切}\theta\in\Theta\text{和}T\in U,$$

则称T_0是θ的一致最小方差无偏估计（记为$UMVUE$）。

首先，我们给出下面重要的Lehmann-Scheffé

定理 设X的样本空间和分布族为$(\mathscr{X}, \beta_{\mathscr{X}})$，$\{p_\theta, \theta\in\Theta\}$，$g(\theta)$为定义于$\Theta$取值于$R_1$的函数，$t(x)$是$\theta$的充分完全统计量，则

（1）若一个只依赖于$t(x)$的估计量$h(t(x))$为$g(\theta)$的无偏估计，则必定是$g(\theta)$的$UMVUE$；

（2）这样的估计量最多只有一个（在概率为1相等的意义下）；

（3）若$g^*(x)$是$g(\theta)$的任一无偏估计，则必存在(1)中所述的$UMVUE$，若后者的方差处处有限（在Θ内），则它是$g(\theta)$的唯一的$UMVUE$（见〔7〕p95定理2.1.1）。

由Lehmann-Scheffé定理，可知一些常见的无偏估计是$UMVUE$。例如，若$X_1,\cdots,X_n$为iid样本，$X_1\sim N(\mu, \sigma^2)$，则

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

和 $$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

分别为μ和σ^2的$UMVUE$；又如，若$X_1,\cdots,X_n$为iid样本，$X_1\sim p(\lambda)$（Poisson分布），则

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

为λ的 $UMVUE$；又$X_1,\cdots,X_n$ 为iid样本，$X_1\sim U(\theta_1,\theta_2)$（在（$\theta_1$，$\theta_2$）内的均匀分布），则

$$\frac{1}{2}[\max_{1\leqslant i\leqslant n}X_i+\min_{1\leqslant i\leqslant n}X_i]$$

为$EX_1=\dfrac{\theta_1+\theta_2}{2}$的$UMVUE$。

然而，并非一切无偏估计都是$UMVUE$，请看下例：

例1 设$X_1,\cdots,X_n$为 iid 样本，X_1 服从对数正态分布，即$\log X_1$服从正态分布$N(\mu，\sigma^2)$。由此知X_1的密度函数为

$$f(x;\ \theta)=\begin{cases}(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left[-\dfrac{1}{2\sigma^2}(\log x-\mu)^2\right]\dfrac{1}{x}, & 当x>0,\\ 0, & 当x<0。\end{cases}$$

其中 $\theta\overset{\triangle}{=}(\mu,\ \sigma^2)$。

由此不难得到（$X_1,\cdots,X_n$）的联合分布密度，并根据因子分解定理和指数族的完全性定理（见〔7〕p80）可得知

$$t=(t_1,t_2)=\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i,\ \sum_{i=1}^{n}(\log X_i-t_1)^2\right)$$

为完全充分统计量，且t_1与t_2独立，其中

$$t_1 \sim N\left(\mu,\ \frac{\sigma^2}{n}\right),\quad \frac{t_2}{\sigma^2} \sim x^2(n-1).$$

容易算出，EX_1为

$$g(\theta) \xlongequal{\triangle} EX_1 = \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

设我们要估计$g(\theta)$。

因为 $E\overline{\overline{X}} = EX_1 = g(\theta) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$,

所以$\overline{\overline{X}}$是$g(\theta)$的无偏估计，但由于$\overline{\overline{X}}$它不是

$$t = (t_1, t_2) = \left(\frac{1}{n}\sum_{i=1}^{n} \log X_i,\quad \sum_{i=1}^{n} (\log X_i - t_1)^2\right)$$

的函数，故根据上述的 Lehmann—Scheffé 定理，它不可能是$g(\theta)$ 的$UMVUE$。

进一步的讨论

注 用样本均值估计总体均值的优越性（就方差小的标准而言），只是在指数族中才对，本例1说明，在这范围之外（对数正态不属于指数族），则可能存在更好的无偏估计. 换句话说，某一统计量的优越性，不仅取决于这个统计量本身，还要看所抽取的样本分布是怎样的。

注 但是就例1而言，

$$\hat{g}(x) = \exp\{t_1(x)\}\psi(t_2(x))$$

是$g(\theta)$的$UMVUE$，其中

$$\psi(t_2) = \Gamma\left(\frac{n-1}{2}\right)\sum_{k=0}^{\infty}\frac{1}{k!\,\Gamma\left(k+\frac{n-1}{2}\right)}\left(\frac{t_2}{4}\right)^k\left(1-\frac{1}{n}\right)^k.$$

事实上，$\because\quad t_1 \sim N\left(\mu,\ \frac{\sigma^2}{n}\right)$,

$\therefore$ e^{t_1}服从对数正态分布。

由 $EX_1 = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}$

得到 $E(e^{t_1}) = \exp\left\{\mu + \frac{\sigma^2}{2n}\right\}$。

又由于 $\frac{t_2}{\sigma^2} \sim x^2(n-1)$，

$$\therefore \quad E\left(\frac{t_2}{\sigma^2}\right)^k = \int_0^\infty t^k \frac{1}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} t^{\frac{n-1}{2}-1} e^{-\frac{t}{2}} dt$$

$$= \frac{2^k \Gamma\left(k + \frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

$$\therefore \quad E(t_2)^k = 2^k \sigma^{2k} \frac{\Gamma\left(k + \frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}, \quad k = 0, 1, 2, \cdots$$

$$\therefore \quad E\psi(t_2) = \Gamma\left(\frac{n-1}{2}\right) \sum_{k=0}^{\infty} \frac{E(t_2^k)}{k!\,\Gamma\left(k + \frac{n-1}{2}\right)} 4^{-k} \left(1 - \frac{1}{n}\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{2^k \sigma^{2k}}{k!} \frac{1}{4^k} \left(1 - \frac{1}{n}\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\sigma^2}{2}\right)^k \left(1 - \frac{1}{n}\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{\left[\frac{\sigma^2}{2}\left(1 - \frac{1}{n}\right)\right]^k}{k!}$$

$$= \exp\left(\frac{\sigma^2}{2}\left(1 - \frac{1}{n}\right)\right)$$

故 $E\hat{g}(x)=E(e^{t_1})E[\psi(t_2)]$

$$=\exp\left(\mu+\frac{\sigma^2}{2n}\right)\text{epx}\left[\frac{\sigma^2}{2}\left(1-\frac{1}{n}\right)\right]$$

$$=\exp\left(\mu+\frac{\sigma^2}{2}\right)=g(\theta).$$

即 $\hat{g}(x)$ 为 $g(\theta)$ 之无偏估计，由于它是充分完全统计量 $t=(t_1, t_2)$ 的函数，故由Lehmann—Scheffé定理知它是 $g(\theta)$ 的唯一的 *UMVUE*.

9—3 设 $X\sim\{\mathbf{f}(x;\theta),\ \theta\in\theta\Theta\}$，$x_1$，…，$x_n$ 是其iid样本，参数 θ 的无偏估计是 X_1，X_2，…，X_n 的对称函数，但不是 θ 的UMVUE的例子

有关定理

定理1 令

$$U\triangleq\{T:\ E_\theta T=\theta,\ E_\theta T^2<\infty,\ 对一切\theta\in\Theta\},$$

$$U_0\triangleq\{\nu:\ E_\theta\nu=0,\ E_\theta\nu^2<\infty,\ 对一切\theta\in\Theta\}.$$

设 U 是非空集合，则 $T_0\in U$ 是 θ 的UMVUE的充分必要条件是

$$E_\theta(\nu T_0)=0,\ 对一切\theta\in\Theta和\nu\in U_0$$

（见[6]第二册之二分册 $P150$）.

定理2 设 X_1，X_2，…，X_n 是服从分布族 $\{f(x;\theta),\ \theta\in\Theta\}$ 中抽取的iid样本，又 θ 的UMVUE $T(X)$ 存在，则 $T(X)$ 是 X_1，X_2，…，X_n 的对称函数.

证明 ∵ $T(X)$ 是 θ 的UMVUE，∴由定理1有

$$E_\theta(T(X)\nu)=0,$$

即 $$\int\cdots\int \nu T(x_1,x_2,\cdots,x_i,\cdots,x_j,\cdots x_n)\prod_{i=1}^{n}f(x_i;\theta)dx_1\cdots dx_n=0.$$

$\because X_1, X_2, \cdots, X_n$是iid的，$\therefore X_i$之密度函数$f(x_i;\theta)$与$X_j$的密度函数$f(x_j;\theta)$形式完全一样，$\therefore$在上式积分号下作变换，以$x_i$代$x_j$，以$x_j$代$x_i$得

$$\int\cdots\int \nu T(x_1,x_2,\cdots,x_j,\cdots,x_i,\cdots,x_n)\prod_{i=1}^{n}f(x_i;\theta)dx_1,\cdots, dx_n=0,$$

因而 $E\nu T(X_1,\cdots,X_i,\cdots,X_j,\cdots,X_n)=E\nu T(X_1,\cdots,X_j,\cdots,X_i,\cdots,X_n)=0.$

$\because$ $T(X_1, X_2, \cdots, X_i, \cdots,X_j,\cdots,X_n)$是$\theta$的UMVUE，

$\therefore$ $T(X_1,X_2,\cdots,X_j,\cdots,X_i,\cdots,X_n)$也是$\theta$的UMVUE。

于是根据UMVUE的唯一性有

$$T(X_1,X_2,\cdots,X_i,\cdots,X_j,\cdots,X_n)=T(X_1,X_2,\cdots,X_j,\cdots,X_i,\cdots,X_n)\ a,s,$$

而$T(X_1, \cdots, X_n)$是$X_1, \cdots, X_n$的确定的函数。$\therefore T(X_1,\cdots,X_n)$是$X_1, \cdots, X_n$的对称函数。

但定理2的逆不真，我们有如下的

反例

设$X_1, X_2, \cdots, X_n$ iid，

$$x_1\sim P(\lambda)=\frac{\lambda^k}{k!}e^{-\lambda}, \quad \lambda>0,$$

则 $$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i, \qquad S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2.$$

都是λ的无偏估计，且它们均是$X_1, X_2, \cdots, X_n$的对称函数。

然而，$\bar{X}$是λ的UMVUE，而S_*^{*2}却不是λ的UMVUE。

(*) 9—4 无偏估计存在，而UMVUE不存在的例子

我们已经知道，无偏估计可以不存在。然而，无偏估计即使存在，而UMVUE也可以不存在，请看下例：

例1 设$\Theta=\{0, 1\}$，$f_\theta(x)$定义为

$$f_0(x)=I_{(0,1)}(x) \qquad ((0, 1)\text{上的均匀分布}),$$

$$f_1(x)=\frac{1}{4}x^{-\frac{3}{4}}I_{(0,1)}(x), \qquad d\mu=dx.$$

任取$\delta\in(0, 1)$，定义估计量$\hat{\theta}_\delta(x)$如下（样本大小为1，样本就记为x）：

$$\hat{\theta}_\delta(x)=h[I_{(0,\delta)}(x)-I_{(1-\delta,1)}(x)],$$

其中 $h=[\delta^{\frac{1}{4}}+(1-\delta)^{\frac{1}{4}}-1]^{-1}$.

则易证对任何的$\delta\in(0, 1)$，$\hat{\theta}_\delta(x)$为θ的无偏估计，且

$$D_{\theta=1}\hat{\theta}_\delta(X)=O(\delta^{\frac{1}{2}})\longrightarrow 0 \quad (\text{取}\delta>0\text{任意小})$$

即存在θ的无偏估计$\hat{\theta}_\delta(x)$，其当$\theta=1$时的方差可小于任何指定值。由此可知，若θ的UMVUE存在（记为$\theta^*(x)$），则必有

$$D_{\theta=1}\theta^*(x)=0.$$

而这只有在

$$P_\theta(\theta^*(X)=1)=1$$

时才有可能。若后一事实成立，将有

$$E_{\theta=0}\theta^*(X)=1.$$

这与θ^*为θ的无偏估计矛盾。这说明θ的UMVUE不存在。

9—5 $\hat{\theta}$是θ的无偏估计，但$g(\hat{\theta})$不是$g(\theta)$的无偏估计的例子（其中$g(X)$是$X=(X_1,\cdots,X_n)$的Borel 可测函数）.

有关定理

显然，如下结论不证自明：

定理 若$g(\hat{\theta})$是$g(\theta)$的无偏估计，则$\hat{\theta}$是θ的无偏估计其中$g(X)$是$\boldsymbol{X}=(X_1, \cdots, X_n)$的任一Borel可测函数.

但其逆不成立，请看

例1 若

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

是均值μ的无偏估计，即$E\overline{X}=\mu$，则当$DX\neq 0$时（此处X是随机变量），$\overline{X}^2$不是μ^2的无偏估计. 事实上

$$E\overline{X}^2=D\overline{X}+(E\overline{X})^2=\frac{DX}{n}+\mu^2\neq\mu^2.$$

例2 若$\boldsymbol{X}\sim N(\mu, \sigma^2)$，则

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

是σ^2的无偏估计.

但S_n^*却不是σ的无偏估计. 事实上

$$\because \quad \frac{(n-1)S_n^{*2}}{\sigma^2}\sim\chi^2(n-1),$$

$$\therefore \quad E\left(\frac{\sqrt{n-1}S_n^*}{\sigma}\right)=\int_0^\infty\sqrt{x}\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}x^{\frac{n-1}{2}-1}e^{-\frac{x}{2}}dx$$

$$=\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\int_0^{\infty}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}dx$$

$$=\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)$$

$$=\frac{\sqrt{2}\ \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

故 $$E(S_n^*)=\sigma\left[\frac{\sqrt{2}\ \Gamma\left(\frac{n}{2}\right)}{\sqrt{n-1}\ \Gamma\left(\frac{n-1}{2}\right)}\right]\neq\sigma.$$

进一步的讨论

注 若 $\hat{\theta}$ 是θ的无偏估计，又g是θ的线性函数，则有 $g(\hat{\theta})$ 是$g(\theta)$的无偏估计。

注 当$n\to\infty$时，例2的S_n^*将是σ的渐近无偏估计，即

$$E(S_n^*)\longrightarrow\sigma\quad（当n\to\infty时）.$$

事实上，利用Stirling公式，可证

$$\frac{\Gamma(p+h)}{\Gamma(p)}\sim p^h\quad（当p充分大时），$$

于是 $$\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}=\frac{\Gamma\left(\frac{n-1}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sim\left(\frac{n-1}{2}\right)^{\frac{1}{2}}=\sqrt{\frac{n-1}{2}}.$$

（当n充分大时），

故有 $$E(S_n^*)=\sigma\left[\frac{\sqrt{2}}{\sqrt{n-1}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]\approx\sigma.$$

9—6 在一定的优良性准则下，用 $S_{n+1}^2 \triangleq \frac{1}{n+1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 去估计 σ^2 较之用 $S_n^{*2} \triangleq \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 去估计 σ^2 更优的例子

我们知道，当仅考虑无偏性时，用 S_n^{*2} 去估计 σ^2 比之用 S_{n+1}^2 去估计 σ^2 更好. 这是因为

$$ES_n^{*2}=\sigma^2,$$

而
$$ES_{n+1}^2=\frac{n-1}{n+1}\sigma^2\neq\sigma^2,$$

即 S_n^{*2} 是 σ^2 的无偏估计，而 S_{n+1}^2 是 σ^2 的有偏估计. 然而，当仅考虑估计方差的大小时，用 S_{n+1}^2 去估计 σ^2 又比之用 S_n^{*2} 去估计 σ^2 更好，这是因为

$$DS_{n+1}^2=\frac{2(n-1)}{(n+1)^2}\sigma^4<DS_n^{*2}=\frac{2\sigma^4}{n-1}.$$

现在，我们要问应该怎样来衡量 S_n^{*2} 与 S_{n+1}^2 作为 σ^2 的估计哪个更好呢？对此，我们有下面的结论：

定理 在均方误差最小的准则下，S_{n+1}^2 较之 S_n^{*2} 作为 σ^2 的估计更为优良.

证明 因为

$$E(S_n^{*2}-\sigma^2)^2=E(S_n^{*2}-ES_n^{*2})^2$$
$$=DS_n^{*2}=\frac{2\sigma^4}{n-1},$$

而
$$E(S_{n+1}^2-\sigma^2)^2=E(S_{n+1}^2-ES_{n+1}^2+ES_{n+1}^2-\sigma^2)^2$$
$$=DS_{n+1}^2+2E\{(S_{n+1}^2-ES_{n+1}^2)(ES_{n+1}^2-\sigma^2)\}$$
$$+[ES_{n+1}^2-\sigma^2]^2$$

$$= \frac{2(n-1)}{(n+1)^2}\sigma^4 + 2(ES_{n+1}^2 - \sigma^2)(ES_{n+1}^2 - ES_{n+1}^2)$$

$$+ \left[\frac{n-1}{n+1}\sigma^2 - \sigma^2\right]^2$$

$$= \frac{2(n-1)\sigma^4}{(n+1)^2} + \frac{4\sigma^4}{(n+1)^2} = \frac{2n+2}{(n+1)^2}\sigma^4 = \frac{2\sigma^4}{n+1},$$

于是 $\quad E(S_{n+1}^2 - \sigma^2)^2 < E(S_n^{*2} - \sigma^2)^2$.

即在均方误差最小的准则下，参数 σ^2 的估计量取 S_{n+1}^2 较之取 S_n^{*2} 更优。这个准则是在兼顾了方差及偏差的绝对值都要求相对地小的情况下确定的。

注 请与例9—11的容许性问题联系起来阅读。

9—7 无偏估计不是一致(相合)估计的例子

定义 设X_1，…，X_n是抽自分布族$\{F(x;\theta),\ \theta\in Ⓗ\}$的子样，$T_n = T_n(X_1, \cdots, X_n)$是$\theta$的估计，如果序列$\{T_n\}$随机收敛到真参数值$\theta$，即对任意$\varepsilon > 0$，

$$\lim_{n\to\infty} P_\theta\{|T_n - \theta| \geqslant \varepsilon\} = 0，\ 对一切\theta\in Ⓗ.$$

则称T_n是θ的一致估计。

从大样本理论的观点看，一致性是对估计量最起码的要求。这个要求也是最容易满足的要求。这主要是由于概率论中的大数定律。然而，在一些情况下，用常见方法构造出来的估计量，也可以没有一致性。下面就是一个例证：

我们知道，若$\boldsymbol{X} \sim U[0,\ \theta]$，即

$$\boldsymbol{X} \sim f(x;\theta) = \begin{cases} \dfrac{1}{\theta} & 当 0 < x \leqslant \theta, \\ 0 & 其他, \end{cases}$$

而X_1，X_2，…，X_n是从$\boldsymbol{X}$中抽取的iid样本，则

$$T_1(X) \triangleq \frac{n+1}{n} \max_{1\leqslant i\leqslant n} X_i$$

和 $$T_2(X_1) = 2X_1$$

都是θ的无偏估计．事实上，因为

$$\max_{1\leqslant i\leqslant n} X_i \sim \begin{cases} \dfrac{nt^{n-1}}{\theta^n}, & \text{当}0<t\leqslant\theta, \\ 0, & \text{其他}. \end{cases}$$

故 $$ET_1(X) = \frac{n+1}{n} E(\max_{1\leqslant i\leqslant n} X_i) = \frac{n+1}{n}\int_0^\theta t\frac{nt^{n-1}}{\theta^n}dt$$

$$= \frac{n+1}{n} \frac{n}{\theta^n} \frac{t^{n+1}}{n+1}\Big|_0^\theta = \theta.$$

$$ET_2(X_1) = 2EX_1 = 2\int_0^\theta t\frac{1}{\theta}dt = \theta,$$

下面，我们来证明$T_1(X)$是θ的一致估计，而$T_2(X_1)$却不是θ的一致估计：

∵ $$E(\max_{1\leqslant i\leqslant n} X_i) = \frac{n\theta}{n+1},$$

$$E[(\max_{1\leqslant i\leqslant n} X_i)^2] = \int_0^\theta t^2\frac{nt^{n-1}}{\theta^n}dt = \frac{n\theta^2}{n+2},$$

故 $$D(\max_{1\leqslant i\leqslant n} X_i) = \frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1}\right)^2 = \frac{n\theta^2}{(n+2)(n+1)^2}.$$

于是，由切比雪夫不等式有

$$P\{|T_1(X)-\theta|\geqslant\varepsilon\} = P\left\{\left|\frac{n+1}{n}\max_{1\leqslant i\leqslant n} X_i - \theta\right|\geqslant\varepsilon\right\}$$

$$\leqslant \frac{D\left(\dfrac{n+1}{n}\max\limits_{1\leqslant i\leqslant n} X_i\right)}{\varepsilon^2} = \left(\frac{n+1}{n}\right)^2 \frac{n\theta^2}{\varepsilon^2(n+2)(n+1)^2}$$

$$\longrightarrow 0 \qquad (n\to\infty).$$

即$T_1(X)$是θ的一致估计．然而

$$P\{|T_2(X_1)-\theta|\geqslant\varepsilon\}=P\{|2X_1-\theta|\geqslant\varepsilon\}$$

$$=\int_{|2t-\theta|\geqslant\varepsilon}\frac{1}{\theta}dt=\frac{1}{\theta}\left[\int_{\frac{1}{2}(\theta+\varepsilon)}^{\theta}dt+\int_0^{\frac{1}{2}(\theta-\varepsilon)}dt\right]$$

$$=\frac{1}{\theta}(\theta-\varepsilon)\not\longrightarrow 0\ (n\to\infty).$$

即$T_2(X_1)$虽是θ的无偏估计，但却不是θ的一致估计。

注 θ的有偏估计

$$T_3(X)=\max_{1\leqslant i\leqslant n}X_i$$

也是θ的一致估计。事实上，

$$P\{|T_3(X)-\theta|\geqslant\varepsilon\}=P\{|\max_{1\leqslant i\leqslant n}X_i-\theta|\geqslant\varepsilon\}$$

$$\leqslant\frac{E(\max\limits_{1\leqslant i\leqslant n}X_i-\theta)^2}{\varepsilon^2}$$

$$=\frac{E\left[\left(\max\limits_{1\leqslant i\leqslant n}X_i-\frac{n\theta}{n+1}\right)-\left(\theta-\frac{n\theta}{n+1}\right)\right]^2}{\varepsilon^2}$$

$$=\frac{D\left(\max\limits_{1\leqslant i\leqslant n}X_i\right)+\left(\theta-\frac{n\theta}{n+1}\right)^2}{\varepsilon^2}$$

$$=\frac{\frac{n\theta^2}{(n+2)(n+1)^2}+\left(\theta-\frac{n\theta}{n+1}\right)^2}{\varepsilon^2}\longrightarrow 0\ (n\to\infty).$$

注 θ的有偏估计$T_4(X_1)=X_1$却不是θ的一致估计。事实上

$$P\{|T_4(X_1)-\theta|\geqslant\varepsilon\}=P\{|X_1-\theta|\geqslant\varepsilon\}$$

$$=\int_{|t-\theta|\geqslant\varepsilon}\frac{1}{\theta}dt=\frac{1}{\theta}\int_0^{\theta-\varepsilon}dt=\frac{\theta-\varepsilon}{\theta}\not\longrightarrow 0\ (n\to\infty).$$

综上所述，可知在一般情况下，无偏性和一致性之间没有一定的关系。

9—8 无偏估计的方差低于Rao-cramér不等式下界的例子

有关定理.

首先让我们叙述Rao-cramér不等式定理：

定理（Rao-cramér不等式）设Θ是实数轴上的一个开区间，$\{f(x;\theta),\theta\in\Theta\}$ 是母体X的分布密度族；$X=(X_1\cdots, X_n)$是从该母体中抽取的一个简单随机样本；$T(X)$是$g(\theta)$的一个无偏估计，且满足条件：

(1) 集合 $S_\theta \overset{\triangle}{=} \{x:f(x;\theta)>0\}$ 与θ无关；

(2) $g'(\theta)$和$\dfrac{\partial f(x;\theta)}{\partial\theta}$存在，且

$$\frac{\partial}{\partial\theta}\int f(x;\theta)dx=\int\frac{2f(x;\theta)}{\partial\theta}dx=0,\ \text{对一切}\theta\in\Theta.$$

$$\frac{\partial}{\partial\theta}\int T(x)L(x;\theta)dx=\int T(x)\ \frac{\partial L(x;\theta)}{\partial\theta}dx,\ \text{对一}$$

切$\theta\in\Theta$.

其中 $L(x;\theta)=\prod\limits_{i=1}^{n}f(x_i;\theta)$

是子样$X=(X_1,\cdots,X_2)$的联合分布密度。

(3) $I(\theta)\overset{\triangle}{=}E_\theta\left(\dfrac{\partial\ln f(X;\theta)}{\partial\theta}\right)^2>0$,

则 $[g'(\theta)]^2\leqslant D_\theta(T)\cdot nI(\theta)$，对一切$\theta\in\Theta$.

如果上式中等号成立，则存在$C(\theta)\neq0$，使得下式成立：

$$\frac{\partial\ln L(x;\theta)}{\partial\theta}=C(\theta)[T(X)-g(\theta)](a.s).$$

如果$g'(\theta_0)\neq 0$，则有

$$D_{\theta_0}(T)\geqslant\frac{[g'(\theta_0)]^2}{nI(\theta_0)}.$$

特别当$g(\theta)=\theta$时，上式可简化为

$$D_\theta(T)\geqslant\frac{1}{nI(\theta)}$$

（见[6]第二册之二分册P153）。

注 也有的书称Rao-cramér不等式为Fréchet-Rao-cramér不等式。

注 条件(2)、(3)常称为正规性条件。

但是，当 Rao-cramér 不等式定理中的正规性条件不满足时，θ的无偏估计量$T(X)$的方差可以低于Rao-cramér不等式的下界。请看

例1 设$X\sim U[0,\theta]$，则

$$X\sim f(x;\theta)=\begin{cases}\frac{1}{\theta}, & \text{当}0\leqslant x\leqslant\theta,\\ 0, & \text{其他}.\end{cases}$$

且$\Theta=(0,\infty)$，当$\theta\neq x$时，$\frac{\partial f(x;\theta)}{\partial\theta}$存在，但当$\theta=x$时，$\frac{\partial f(x;\theta)}{\partial\theta}$不存在，因此正规性条件不满足。

此时 $\ln f=-\ln\theta$，$\frac{\partial\ln f}{\partial\theta}=-\frac{1}{\theta}$。

我们有 $I_{(\theta)}=E_\theta\left(\frac{\partial\ln f(x;\theta)}{\partial\theta}\right)^2=\int_0^\theta\frac{1}{\theta^3}dx=\frac{1}{\theta^2}$，

∴ Rao-cramér不等式的下界为

$$\frac{1}{nI(\theta)}=\frac{\theta^2}{n}.$$

而 $E\left[\frac{n+1}{n}\max(X_1,\cdots,X_n)\right]=\frac{n+1}{n}\int_0^\theta x\frac{1}{\theta^n}nx^{n-1}dx=\theta$,

即$\frac{n+1}{n}\max(X_1,\cdots,X_n)$是$\theta$的无偏估计。

又$\frac{n+1}{n}\max(X_1\cdots,X_n)$的方差为

$$D\left(\frac{n+1}{n}\max(X_1,\cdots,X_n)\right)$$

$$=\left(\frac{n+1}{n}\right)^2\left\{\int_0^\theta x^2\frac{nx^{n-1}}{\theta^n}dx-\left[E(\max(X_1,\cdots,X_n))\right]^2\right\}$$

$$=\frac{(n+1)^2}{n^2}\left[\frac{n\theta^2}{n+2}-\left(\frac{n\theta}{n+1}\right)^2\right]=\frac{\theta^2}{n(n+2)}<\frac{\theta^2}{n}.$$

即当Rao-cramér 不等式定理的正规性条件不成立时，θ的无偏估计的方差可以低于Rao-cramér不等式的下界。

9—9 UMVUE其方差达不到Rao-cramér不等式下界的例子

有关定理

首先，不难证明如下结论：

定理(在一定条件下，见例9-8之Rao-cramér 不等式定理)方差达到Rao-cramér不等式下界的无偏估计，必定是UMVUE。

但定理的逆不真，请看

例1 设母体X～$N(\mu,\sigma^2)$，其中μ和σ^2是未知参数．设$X=(X_1,\cdots,X_n)$是从该母体中抽取的iid样本。则不难知道

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^n(X_i-\overline{X})^2$$

是σ^2的UMVUE。但由

$$\frac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$$

知 $D\left(\frac{(n-1)S_n^{*2}}{\sigma^2}\right) = 2(n-1)$，

从而 $D\left(S_n^{*2}\right) = \frac{2\sigma^4}{n-1}$。

但 $I(\sigma^2) = E_{\sigma^2}\left(\frac{\partial \ln f}{\partial \sigma^2}\right)^2$

$$= \int_{-\infty}^{\infty}\left[\frac{(x-\mu)^2}{2\sigma^4} - \frac{1}{2\sigma^2}\right]^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2\sigma^4},$$

其中 $f = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

故Rao-cramér不等式的下界为

$$\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n},$$

$$\therefore \quad D(S_n^{*2}) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n}.$$

即σ^2的UMVUE S_n^{*2}其方差$D(S_n^{*2})$达不到C—R不等式的下界。

9—10 充分统计量不是有效估计量的例子

有关定义和定理

定义 如果T是参数θ的一个无偏估计，它的方差达到Rao-cramér不等式的下界，那么称T是θ的有效估计量。

我们有如下正面的命题

定理 若$T(X)$是θ的无偏估计量，则当$T(X)$是θ的有效估计量时，必有$T(X)$是θ的充分统计量（见[6]第二册之二分册P166定理4）。

但其逆不成立，请看

例1 设$X \sim N(\mu,\sigma^2)$，μ已知，则

$$S_n^{*^2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

是σ^2的充分统计量，但

$$D\left(S_n^{*^2}\right)=\frac{2\sigma^4}{n-1}>\frac{2\sigma^4}{n},$$

其中$\frac{2\sigma^4}{n}$是$C—R$不等式的下界。

故$S_n^{*^2}$不是σ^2的有效估计量。

进一步的讨论

注 设$T(X)$是θ的无偏估计，则当$T(X)$是θ的充分统计量且满足

$$\frac{\partial \ln f_T(t,\theta)}{\partial\theta}=C(\theta)(T-\theta)\quad (a.S)$$

（其中$f_T(t,\theta)$为T的分布密度，$C(\theta)\neq 0$）时，$T(X)$也是θ的有效估计量。

证明 记$f(x|t)$是$X=(X_1\cdots,X_n)$关于$T(X)=t$的条件密度，由于$T(X)$是θ的充分统计量，所以它与θ无关，且有

$$f(x,\theta)=f(x|t)f_T(t,\theta),$$

其中$f(x,\theta)$是$X=(X_1,\cdots,X_n)$的联合分布密度，$f_T(t,\theta)$是T的分布密度。

$$\therefore\qquad \frac{\partial \ln L(x;\theta)}{\partial\theta}=\frac{\partial \ln f_T(t,\theta)}{\partial\theta}=C(\theta)(T-\theta),$$

于是 $$E\left(\frac{\partial \ln L}{\partial\theta}\right)^2=E\left(\frac{\partial \ln f_T(t,\theta)}{\partial\theta}\right)^2=C^2(\theta)E(T-\theta)^2$$

$$=C^2(\theta)E(T-ET)^2=C^2(\theta)DT,$$

∴ $$1=\frac{\partial\theta}{\partial\theta}=\frac{\partial}{\partial\theta}ET=\frac{\partial}{\partial\theta}\int t f_T(t,\theta)dt$$

$$=\int t\frac{\partial}{\partial\theta}f_T(t,\theta)dt=\int(t-\theta)\frac{\partial f_T(t,\theta)}{\partial\theta}dt$$

$$=\int(t-\theta)\frac{\partial \ln f_T(t,\theta)}{\partial\theta}f_T(t,\theta)dt$$

$$=\int(t-\theta)C(\theta)(t-\theta)f_T(t,\theta)dt$$

$$=c(\theta)\int(t-\theta)^2 f_T(t,\theta)dt=c(\theta)DT.$$

∵ $$E\frac{\partial \ln L}{\partial\theta}=\sum_{i=1}^{n}E\ \frac{\partial \ln f(x_i;\theta)}{\partial\theta}=\sum_{i=1}^{n}\int\frac{\partial \ln f}{\partial\theta}fdx$$

$$=\sum_{i=1}^{n}\int\frac{\partial \ln f}{\partial f}\frac{\partial f}{\partial\theta}fdx=\sum_{i=1}^{n}\int\frac{\partial f}{\partial\theta}dx$$

$$=\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\int fdx=0,$$

∴ $$D\left(\frac{\partial \ln L(x;\theta)}{\partial\theta}\right)=E\left(\frac{\partial \ln L(x;\theta)}{\partial\theta}\right)^2$$

$$=E\left(\frac{\partial \ln \prod_{i=1}^{n}f(x_i;\theta)}{\partial\theta}\right)^2=\sum_{i=1}^{n}\left(\frac{\partial \ln f}{\partial\theta}\right)^2$$

$$=\sum_{i=1}^{n}I(\theta)=nI(\theta).$$

故 $$nI(\theta)=c^2(\theta)DT=\frac{C^2(\theta)D^2T}{DT}=\frac{1}{DT}.$$

即 $$DT=\frac{1}{nI(\theta)},$$

∴ T是θ的有效估计量.

(*)9—11 参数的UMVUE不是参数的可容许估计的例子

定义 设d为θ的一个估计，若不存在另一个估计d^*，使

$$R(\theta,d^*)\leqslant R(\theta,d) \quad 对一切\theta\in\Theta成立$$

且不等号至少对一个$\theta\in\Theta$成立，则称d为可容许估计. 反之，则称d是不可容许的. 其中

$$R(\theta,d)=EL(\theta,d)$$

为估计d的风险函数，$L(\theta,d)$为估计d之损失函数.

在估计问题中，在一定的最优准则下得到的参数θ的估计量，有不少是可容许的. 但是，也存在着这样的情况：估计量即使是θ的UMVUE,但它却是不可容许的. 我们有下面的例子.

例1 设$X_1,X_2,\cdots,X_n$是iid样本，$X_1\sim N(\mu,\sigma^2)$，$\theta=(\mu,\sigma^2)$，欲估计σ^2. 定义损失函数

$$L(\sigma^2,d)=\frac{(\sigma^2-d)^2}{\sigma^4}.$$

我们知道，常用的估计量

$$s_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\triangleq\frac{1}{n-1}s^2$$

是σ^2的 UMVUE. 然而它是不可容许的. 为了证明这一点，考虑形如cs^2的估计，c为待定常数. 注意到

$$\frac{s^2}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2\sim x^2(n-1)$$

有 $$E\left(\frac{s^2}{\sigma^2}\right)=n-1,\quad D\left(\frac{s^2}{\sigma^2}\right)=2(n-1),$$

∴ $$R(\sigma^2,cs^2)=EL(\sigma^2,cs^2)=E\left[\frac{(cs^2-\sigma^2)^2}{\sigma^4}\right]$$

$$=c^2E\left(\frac{s^2}{\sigma^2}\right)^2-2cE\left(\frac{s^2}{\sigma^2}\right)+1$$

$$=c^2\left[D\left(\frac{s^2}{\sigma^2}\right)+E^2\left(\frac{s^2}{\sigma^2}\right)\right]-2cE\left(\frac{s^2}{\sigma^2}\right)+1$$

$$=2(n-1)c^2+[1-(n-1)c]^2.$$

其最小值在$c=\dfrac{1}{n+1}$处达到。

估计量s_n^{*2}的风险为

$$R(\sigma^2,s_n^{*2})=EL(\sigma^2,s_n^{*2})=E\frac{(\sigma^2-s_n^{*2})^2}{\sigma^4}$$

$$=E\left(\frac{s_n^{*4}-2\sigma^2s_n^{*2}+\sigma^4}{\sigma^4}\right)=E\left(\frac{s_n^{*2}}{\sigma^2}\right)^2-2E\left(\frac{s_n^{*2}}{\sigma^2}\right)+1$$

$$=\left[D\left(\frac{s_n^{*2}}{\sigma^2}\right)+E^2\left(\frac{s_n^{*2}}{\sigma^2}\right)\right]-2E\left(\frac{s_n^{*2}}{\sigma^2}\right)+1$$

$$=\left[\frac{2}{n-1}+1\right]-2+1=\frac{2}{n-1}.$$

估计量 $$\frac{n-1}{n+1}s_n^{*2}=\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

之风险为$R\left(\sigma^2,\dfrac{n-1}{n+1}s_n^{*2}\right)=R\left(\sigma^2,\dfrac{1}{n+1}s^2\right)$

$$=2(n-1)\left(\frac{1}{n+1}\right)^2+\left[1-(n-1)\frac{1}{n+1}\right]^2$$

$$=\frac{2}{n+1}.$$

显然有 $$R\left(\sigma^2,\frac{n-1}{n+1}s_n^{*2}\right)<R(\sigma^2,s_n^{*2}),$$

故s_n^{**}是不可容许的（请与例9—6联系起来阅读）。

下面我们再给一个著名例子，它是stein给出的。

例2 设X_1，X_2，…，X_n为iid样本，X_1服从k维正态分布$N(\theta,I_k)$。欲估计$\theta=(\theta_1,\cdots,\theta_k)^\tau$，其中记号$I_k$表示$k$阶单位阵，$(\theta_1,\cdots,\theta_k)^\tau$表示$\begin{pmatrix}\theta_1\\ \vdots\\ \theta_k\end{pmatrix}$之转置。定义损失函数为$L(\theta,d)=\|\theta-d\|^2$。通常的估计量是

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i.$$

它是θ的UMVUE。但是stein在1956年证明了当参数$k\geqslant3$时，这个估计量是不可容许的：估计量

$$\theta^*=\left(1-\frac{k-2}{n\|\bar{X}\|^2}\right)\bar{X} \tag{1}$$

在上述损失函数下，一致地优于$\bar{X}$。

下面我们来证明这个结论。由于stein 1956年的证明过分冗长，这里我们采用Hudon 1978年给出的较为简练的分部积分法来证明。

$$\theta^*=\left(1-\frac{k-2}{n\|\bar{X}\|^2}\right)\bar{X}=\bar{X}-\frac{k-2}{n\bar{X}^\tau\bar{X}}\bar{X}$$

的风险函数为

$$R(\theta,\theta^*)=EL(\theta,\theta^*)=E\left\|\bar{X}-\frac{k-2}{n\bar{X}^\tau\bar{X}}\bar{X}-\theta\right\|^2$$

$$=E\left[(\bar{X}-\theta)^\tau-\frac{(k-2)\bar{X}^\tau}{n\bar{X}^\tau\bar{X}}\right]\left[(\bar{X}-\theta)-\frac{(k-2)\bar{X}}{n\bar{X}^2\bar{X}}\right]$$

$$=E\|\bar{X}-\theta\|^2+\frac{(k-2)^2}{n^2}E\left(\frac{1}{\bar{X}^\tau\bar{X}}\right)$$

$$-\frac{2(k-2)}{n}E\left[\frac{(\bar{X}-\theta)^{\tau}\bar{X}}{\bar{X}^{\tau}\bar{X}}\right]. \qquad (*)$$

注意到 $\bar{X}\sim N\left(\theta,\ \frac{1}{n}I_k\right)$,

知(*)式其中的

$$E\left[(\bar{X}-\theta)^{\tau}\frac{\bar{X}}{\bar{X}^{\tau}\bar{X}}\right]$$

$$=\int_{R_k}\cdots\int\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k\frac{\sum_{i=1}^{k}y_i(y_i-\theta)}{\sum_{j=1}^{k}y_j^2}\exp$$

$$\times\left\{-\frac{n\sum_{j=1}^{k}(y_j-\theta_j)^2}{2}\right\}\prod_{j=1}^{k}dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k\sum_{i=1}^{k}\int_{R_k}\cdots\int\frac{y_i(y_i-\theta_i)}{\sum_{j=1}^{k}y_j^2}\exp$$

$$\times\left\{-\frac{n\sum_{j=1}^{k}(y_j-\theta_j)^2}{2}\right\}\prod_{j=1}^{k}dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k\sum_{i=1}^{k}\int_{-\infty}^{\infty}\frac{y_i}{\sum_{j=1}^{k}y_j^2}(y_i-\theta_i)e^{-\frac{n(y_i-\theta_i)^2}{2}}$$

$$\times dy_i\int_{R_{k-1}}\cdots\int\exp\left\{-\frac{\sum_{j=1,j\neq i}^{k}(y_j-\theta_j)^2}{2}\right\}\prod_{j=1,j\neq i}^{k}dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k \sum_{i=1}^{k}\int_{-\infty}^{\infty}\frac{y_i}{\sum_{j=1}^{k} y_j^2} d\left(-\frac{1}{n}e^{-\frac{n(y_i-\theta_i)^2}{2}}\right)$$

$$\times\int_{R_{k-1}}\cdots\int \exp\left\{-\frac{n}{2}\sum_{\substack{j=1\\ j\neq i}}^{k}(y_j-\theta_j)^2\right\}\prod_{\substack{j=1\\ j\neq i}}^{k} dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k \sum_{i=1}^{k}\int_{-\infty}^{\infty}\frac{1}{n}\exp\left\{-\frac{n(y_i-\theta_i)^2}{2}\right\}$$

$$\times d\left(\frac{y_i}{\sum_{j=1}^{k} y_j^2}\right)\cdot\int_{R_{k-1}}\cdots\int \exp$$

$$\times\left\{-\frac{n}{2}\sum_{\substack{j=1\\ j\neq i}}^{k}(y_j-\theta_j)^2\right\}\prod_{\substack{j=1\\ j\neq i}}^{k} dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k \sum_{i=1}^{k}\int_{-\infty}^{\infty}\frac{\sum_{j=1}^{k} y_j^2-2y_i^2}{n\left(\sum_{j=1}^{k} y_j^2\right)^2}e^{-\frac{n(y_i-\theta_i)^2}{2}}dy_i\,\cdot$$

$$\times\int_{R_{k-1}}\cdots\int \exp\left\{-\frac{n}{2}\sum_{\substack{j=1\\ j\neq i}}^{k}(y_j-\theta_j)^2\right\}\prod_{\substack{j=1\\ j\neq i}}^{k} dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k \sum_{i=1}^{k}\int_{R_k}\cdots\int\frac{\sum_{j=1}^{k} y_j^2-2y_i^2}{n\left(\sum_{j=1}^{k} y_j^2\right)^2}$$

$$\times\exp\left\{-\frac{n\sum_{j=1}^{k}(y_j-\theta_j)^2}{2}\right\}\prod_{j=1}^{k} dy_j$$

$$=\left(\frac{\sqrt{n}}{\sqrt{2\pi}}\right)^k\int_{R_k}\cdots\int\frac{k-2}{n\sum_{j=1}^{k}y_j^2}\exp\left\{-\frac{n\sum_{j=1}^{k}(y_j-\theta_j)^2}{2}\right\}$$

$$\times\prod_{j=1}^{k}dy_j=\frac{k-2}{n}E\left(\frac{1}{\bar{X}'\bar{X}}\right)$$

$$=\frac{k-2}{n}E\left(\frac{1}{\|\bar{X}\|^2}\right).$$

代入(*)式得

$$R(\theta,\theta^*)=E\left\|\bar{X}-\frac{k-2}{n\bar{X}'\bar{X}}\bar{X}-\theta\right\|^2$$

$$=E\|\bar{X}-\theta\|^2-\frac{(k-2)^2}{n^2}E\left(\frac{1}{\|\bar{X}\|^2}\right)$$

$$=R(\theta,\bar{X})-\frac{(k-2)^2}{n^2}E\left(\frac{1}{\|\bar{X}\|^2}\right)<R(\theta,\bar{X})$$

(对一切$\theta\in R_k$)，

∴ $\bar{X}$是不可容许的。

进一步的讨论

注 在例1中，即使是

$$\frac{n-1}{n+1}s_n^{*2}=\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

也是不可容许的。可以证明

$$g^*(x_1,\cdots,x_n)=\min\left\{\frac{1}{n+1}\sum_{i=1}^{n}(x_i-\bar{x})^2,\right.$$

$$\left.\frac{1}{n+2}\sum_{i=1}^{n}x_i^2\right\}\qquad(2)$$

一致地优于$\frac{1}{n+1}\sum_{i=1}^{n}(x_i-\bar{x})^2$。这是stein证明的一个结果，但证明需要用到较多的知识，我们不在此给出了。

注 在例2中的$\bar{X}$既是UMVUE，也是最优不变估计(在平移变换群下）和Minimax估计。因此这个例2告诉我们，不可把一种最优准则绝对化。也应当注意，象(1)、(2)两式给出的估计，目前也没有在应用问题中普遍使用。这除了由于习惯原因外，还应当了解，这些估计的优越性是与一定的损失函数相联系在一起的，不能认为这种损失函数一定符合问题的实际，因此人们宁肯使用直观上看来简单明了的估计S_n^{*2}和$\bar{X}$，而不愿使用较为复杂的估计量(1)和(2)。

9—12 某些条件不满足，但似然方程仍然存在一致(相合)解并且满足渐近正态性的例子

有关定理

关于似然方程解的性质，我们有如下两个定理。

定理1 设母体X具有分布密度族$\{f(x;\theta),\theta\in\Theta\}$，$\Theta$为$R_1$上的一个开集，$X_1, X_2, \cdots, X_n$是抽自母体的$iid$样本，$\frac{\partial f(x;\theta)}{\partial\theta}$在$\Theta$上处处有限，且对任何$\theta_1\neq\theta_2$有不同的分布，又假定$E|l_nf(x;\theta)|<\infty$，对一切$\theta\in\Theta$。则对任何$\theta\in\Theta$，方程

$$\frac{\partial \ln L}{\partial\theta}=\sum_{i=1}^{n}\frac{\partial \ln f(x_i;\theta)}{\partial\theta}=0$$

有一个解$\hat{\theta}_n$，满足

$$P_\theta(\hat{\theta}_n\longrightarrow\theta)=1,$$

即 $\hat{\theta}_n$ 是θ的强一致（相合）解（见〔7〕p204定理2.6.4）。

定理2 设母体$\boldsymbol{X}\sim\{f(x;\theta),\ \theta\in\Theta\}$，$\Theta$为 R_1 上的一个开集，X_1，X_2，…，X_n为抽自母体的iid样本，假定

（1）$\int\frac{\partial f(x;\theta)}{\partial\theta}dx=0,\quad \int\frac{\partial^2 f(x;\theta)}{\partial\theta^2}dx=0$

对一切$\theta\in\Theta$。

（2）$0<I(\theta)=E_J\left(\frac{\partial\ln f(x;\theta)^2}{\partial\theta}\right)^2<\infty$。

（3）$\left|\frac{\partial^3\ln f(x;\theta)}{\partial\theta^3}\right|<H(x)$，且 $EH(X)<M<\infty$。

（4）不同的θ值，对应$\boldsymbol{X}$不同的分布，则对任何$\theta\in\Theta$,似然方程有一个解$\hat{\theta}_n$，满足

$$\sqrt{n}(\hat{\theta}_n-\ell)\xrightarrow{L}N\left(0,\ \frac{1}{I(\theta)}\right).$$

其中L表示依分布收敛（见〔7〕p216定理2.2.6）。

但是，有时会遇到这样的例子，上述定理1，定理2的条件虽不满足，但似然方程仍然存在一致（相合）解$\hat{\theta}_n$，并且满足

$$\sqrt{n}(\hat{\theta}_n-\theta)\xrightarrow{L}N\left(0,\ \frac{1}{I(\theta)}\right).$$

请看下面Kulldorf给出的一个例子。

例1 设X～N（0，θ），$\theta>0$；X_1，X_2，…，X_n是抽自母体X的iid样本。

于是

$$\mathbf{L}(x;\theta)=\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{\sum_{i=1}^{n}x_i^2}{2\theta}},$$

$$\ln L = -\frac{n}{2}\ln 2\pi\theta - \frac{\sum_{i=1}^{n} x_i^2}{2\theta},$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{\sum_{i=1}^{n} x_i^2}{2\theta^2} - \frac{2\pi n}{4\pi\theta},$$

∴ 似然方程

$$\frac{\partial \ln L}{\partial \theta} = 0$$

的解为 $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2$.

且有 $EX^2 = \theta$，$DX^2 = 3\theta^2 - \theta^2 = 2\theta^2$，

又 $f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}}$，

$$\ln f(x;\theta) = -\frac{1}{2}\ln\theta - \frac{1}{2}\ln 2\pi - \frac{x^2}{2\theta},$$

$$\frac{\partial \ln f(x;\theta)}{\partial \theta} = \frac{x^2}{2\theta^2} - \frac{1}{2\theta},$$

∴ $I(\theta) = E_\theta\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2$

$$= \int_{-\infty}^{\infty}\left(\frac{x^2}{2\theta^2} - \frac{1}{2\theta}\right)^2 \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}} dx$$

$$= \frac{3\theta^2}{4\theta^4} - \frac{\theta}{2\theta^3} + \frac{1}{4\theta^2} = \frac{1}{2\theta^2}.$$

由柯尔莫哥洛夫(Колмогоров)强大数律，知

$$\hat{\theta}_n \xrightarrow{a.s} \theta,$$

且由中心极限定律

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \theta\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i^2 - \sqrt{n}\,\theta$$

$$= \frac{\sum_{i=1}^{n} X_i^2 - n\theta}{\sqrt{n}} = \sqrt{2}\,\theta\,\frac{\sum_{i=1}^{n} X_i^2 - n\theta}{\sqrt{2n\theta^2}}$$

$$= \sqrt{2}\,\theta\,\frac{\sum_{i=1}^{n} X_i^2 - E\left(\sum_{i=1}^{n} X_i^2\right)}{\sqrt{D\left(\sum_{i=1}^{n} X_i^2\right)}} \xrightarrow{L} N(0,\ 2\theta^2)$$

$$= N\left(0,\ \frac{1}{I(\theta)}\right).$$

然而，$$\frac{\partial^3 \ln f(x;\theta)}{\partial\theta^3} = -\frac{1}{\theta^3} + \frac{3x^2}{\theta^4} \longrightarrow \infty \quad (\text{当}\theta \to 0\text{时}),$$

因此，在 $0<\theta<\infty$ 时，$\frac{\partial^3 \ln f(x;\theta)}{\partial\theta^3}$ 不是有界的，即上述定理2的条件(3) 不满足（此例可见*G. kulldorff*, *On the Conditions for Consistency and Asymptotic Efficiency of Maximum Likelihood Estimates. Skand Aktuarietidskr.* 40(1957) *p*129—144)。

(*)9—13 参数θ的一致(相合)渐近正态估计$\hat{g}$在θ点的渐近方差$\frac{v(\theta)}{n}$中的$v(\theta)$低于C-R下界的例子

有关定义

定义1 设$\hat{g}(X_1, \cdots, X_n)$为$g(\theta)$的估计，若存在$v(\theta)$，$0 \leqslant v(\theta) < \infty$，对任何$\theta \in \Theta$，使得

$$\sqrt{n}\left[\hat{g}(X_1, X_2, \cdots, X_n) - g(\theta)\right] \xrightarrow{L} N(0, v(\theta))$$ 对

一切$\theta \in \Theta$则称 $\hat{g}$为$g(\theta)$的相合渐近正态估计 (*Consistent Asymptotically Normal*)简称CAN估计.换句话说CAN估计是既是相合,其分布又渐近于正态分布的那种估计.$\frac{v(\theta)}{n}$有时称为$\hat{g}$在θ点的渐近方差.

从定义1中不难知道，θ的CAN估计$\hat{g}$必是θ的弱相合估计(即一致性估计).

在较早的著作中，曾引进“有效估计”的概念. 这是基于C—R不等式而来的. 设$\hat{g}$为$g(\theta)$的无偏估计，则由C—R不等式知（在一定的正规性条件下）

$$D(\hat{g}) \geqslant \frac{[g'(\theta)]^2}{nI(\theta)}.$$

因此，把二者的比值（$D(\hat{g})$作分母）称作$\hat{g}$的效率，当效率为1时，称$\hat{g}$为$g(\theta)$的有效估计.

这个定义的缺点是明显的.因为能达到C—R不等式下界的情况不是常规，而是一种例外. 这样，在大多数情况下，有效估计将不存在.另外,C—R不等式的成立有一定的条件，当这些

条件不成立时，$C—R$不等式可以不对(见例9—8)，这时，依据它所提供的下限去定义效率就不合理了. 因此，有人考虑将"有效"的要求放宽为所谓"渐近有效"，即当$n\to\infty$时，上述比值趋于1. 这个定义虽然部分地解决了上述问题，但仍存在重要的缺点，即方差$D(\hat{g})$可能不存在或难于确定. 因此，在一般情况下，很难验证$\hat{g}$是否具有这个性质.

然而，CAN估计却是常见的，在很一般的条件下，矩估计是CAN估计，而在较有限制性的条件下，似然方程有一根为CAN估计. 自然地，若$\hat{g}_1$和$\hat{g}_2$都是$g(\theta)$的CAN估计，即

$$\sqrt{n}\left[\hat{g}_i(X_1,\ \cdots,\ X_n)-g(\theta)\right]\xrightarrow{L}N\ (0,\ \ v_i(\theta)),$$

$i=1,2$. 若$\hat{g}_1$的渐近方差$v_1(\theta)<v_2(\theta)$($\hat{g}_2$的渐近方差)，则我们认为$\hat{g}_1$渐近地优于$\hat{g}_2$. 这可以作如下的解释：任给$c>0$，问"$\hat{g}_i$与$g(\theta)$的偏差不超过$\frac{c}{\sqrt{n}}$"的概率有多大？由定义可知，近似地有

$$p_\theta\left\{\left|\hat{g}_i-g(\theta)\right|\leqslant\frac{c}{\sqrt{n}}\right\}\approx\frac{1}{\sqrt{2\pi}}\int_{-\frac{c}{v_i(\theta)}}^{\frac{c}{v_i(\theta)}}e^{-\frac{x^2}{2}}dx.$$

由此可知，若

$$v_i(\theta)<v_2(\theta),$$

则当n充分大时，有

$$p_\theta\left\{\left|\hat{g}_1-g(\theta)\right|\leqslant\frac{c}{\sqrt{n}}\right\}>p_\theta\left\{\left|\hat{g}_2-g(\theta)\right|\leqslant\frac{c}{\sqrt{n}}\right\}$$

对一切$c>0$成立.

就是说，当样本大小n足够大时，使用$\hat{g}_1$去估计$g(\theta)$较之使用$\hat{g}_2$去估计$g(\theta)$，更可能得到较准确的估计，正是在这个意义上，我们说$\hat{g}_1$渐近地优于$\hat{g}_2$.

因此，自然就提出了找渐近方差最小的CAN估计的问题（这种估计为最优渐近正态估计）。然而，在一切CAN估计中，$v(\theta)$的下确界如何呢？

Fisher 早就猜测这个下确界就是Fisher信息量$I(\theta)$的倒数$\frac{1}{I(\theta)}$。Fisher 是从他对极大似然估计的研究作出这一猜测的（他认为极大似然估计是最好的估计，其渐近方差应是最小的）。现在我们从 $C—R$ 不等式的角度观察，也觉得这个猜测是可信的。这导致了如下的定义。

定义2 设对任何$\theta\in\Theta$，当$n\to\infty$时，

$$\sqrt{n}[\hat{g}(X_1,X_2,\cdots,X_n)-g(\theta)]\xrightarrow{L}N\left(0,\frac{[g'(\theta)]^2}{I(\theta)}\right),$$

则称$\hat{g}$为$g(\theta)$的最优渐近正态估计（*Best asymptotically normal estimate*），简称BAN估计（此处假设X_1，X_2，$\cdots$，X_n为iid样本）。当$g(\theta)=\theta$时，

$$\frac{[g'(\theta)]^2}{I(\theta)}=\frac{1}{I(\theta)}.$$

然而，在五十年代，Hodges举出一个反例，说明$v(\theta)$可以低于$C—R$下界，这样估计叫"超有效估计"（"*Super-efficient estimate*"）。这样一来，就打破了Fisher的这个猜测。这一事实引起了一些学者的兴趣。经过一些学者研究，证明了Fisher的猜测在一定的限制性条件下是对的（具体地说要排除某些象Hodges例子中的"病态性"的CAN估计，只对剩下的CAN估计求渐近方差的下确界），而且这种所谓的"超有效性"只能出现在很稀有的、非一致收敛的情况中。

下面我们来给出Hodges的例子，并给予详细的证明。

例1 设$\hat{\theta}$为θ的CAN估计，且参数空间Θ包含$\theta=0$的一个邻域，记

$$\sqrt{n}\,[\hat{g}(X_1,\ \cdots,\ X_n)-\theta]\xrightarrow{L}N(0,\ v(\theta)).$$

定义 $g^*(X_1,\cdots,X_n)$

$$=\begin{cases}a\hat{g}(X_1,X_2,\cdots,X_n), & \text{当}\ |\hat{g}(X_1,\cdots,X_n)|<n^{-\frac{1}{4}},\\ \hat{g}(X_1,X_2,\cdots,X_n), & \text{当}\ |\hat{g}(X_1,\cdots,X_n)|\geqslant n^{-\frac{1}{4}}.\end{cases}$$

则当$\theta=0$时，

$$\lim_{n\to\infty}p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=\lim_{n\to\infty}p_\theta\left\{\left|\sqrt{n}\,\hat{g}\right|<n^{\frac{1}{4}}\right\}.$$

∵ $\theta=0$时，$\sqrt{n}(\hat{g}-\theta)=\sqrt{n}\,\hat{g}$的极限分布为正态分布，而$n^{\frac{1}{4}}\longrightarrow\infty(n\to\infty)$，$-n^{\frac{1}{4}}\longrightarrow-\infty\ (n\to\infty)$

∴ 当$\theta=0$时，

$$\lim_{n\to\infty}p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=\lim_{n\to\infty}p_\theta\left\{\left|\sqrt{n}\,\hat{g}\right|<n^{\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{-n^{\frac{1}{4}}<\sqrt{n}\,\hat{g}<n^{\frac{1}{4}}\right\}=1.$$

又当$\theta>0$时，

$$\lim_{n\to\infty}p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=\lim_{n\to\infty}p_\theta\left\{-n^{-\frac{1}{4}}<\hat{g}<n^{-\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{-n^{-\frac{1}{4}}-\theta<\hat{g}-\theta<n^{-\frac{1}{4}}-\theta\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}\left(-n^{-\frac{1}{4}}-\theta\right)<\sqrt{n}(\hat{g}-\theta)<\sqrt{n}\left(n^{-\frac{1}{4}}-\theta\right)\right\}.$$

∵ $\sqrt{n}(\hat{g}-\theta)$的极限分布为正态分布，而$\theta>0$，

∴ 当$n\to\infty$时，$\sqrt{n}\left(-n^{-\frac{1}{4}}-\theta\right)$与$\sqrt{n}\left(n^{-\frac{1}{4}}-\theta\right)$都$\to-\infty$。

故当$\theta>0$时，有

$$\lim_{n\to\infty}p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=0.$$

同理可证，当$\theta<0$时，也有

$$\lim_{n\to\infty} p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=0$$

综合得当$\theta\neq 0$时，有

$$\lim_{n\to\infty} p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=0。$$

于是有

$$\lim_{n\to\infty} p_\theta\left\{|\hat{g}(X_1,\cdots,X_n)|<n^{-\frac{1}{4}}\right\}=\begin{cases}1, & \text{当}\theta=0\text{时}, \\ 0, & \text{当}\theta\neq 0\text{时}。\end{cases}$$

由(1) 的结果，可证

$$\sqrt{n}\left[g^*(X_1,\cdots,X_n)-\theta\right]\xrightarrow{L}N(0,\ v^*(\theta))。$$

其中
$$V^*(\theta)=\begin{cases}\alpha^2 v(\theta), & \text{当}\theta=0。\\ v(\theta), & \text{当}\theta\neq 0。\end{cases} \tag{2}$$

事实上，当$\theta\neq 0$时

$$\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(g^*-\theta)<x\right\}$$

$$=\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(\hat{g}-\theta)<x,\quad |\hat{g}|\geqslant n^{-\frac{1}{4}}\right\}$$

$$+\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(\alpha\hat{g}-\theta)<x,\quad |\hat{g}|<n^{-\frac{1}{4}}\right\}。$$

上式右边第二项

$$\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(\alpha\hat{g}-\theta)<x,\quad |\hat{g}|<n^{-\frac{1}{4}}\right\}$$

$$\leqslant\lim_{n\to\infty} p_\theta\left\{|\hat{g}|<n^{-\frac{1}{4}}\right\}=0,$$

∴　当$\theta\neq 0$时，

$$\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(g^*-\theta)<x\right\}$$

$$=\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(\hat{g}-\theta)<x,\quad |\hat{g}|\geqslant n^{-\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty} p_\theta\left\{\sqrt{n}(\hat{g}-\theta)<x\right\}$$

$$-\lim_{n\to\infty}p_\theta\left\{\sqrt{n}(\hat{g}-\theta)<x,|\hat{g}|<n^{-\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}(\hat{g}-\theta)<x\right\}=\int_{-\infty}^{x}N(0,v(\theta))dt,$$

即当$\theta\neq 0$时，有

$$\sqrt{n}(g^*-\theta)\xrightarrow{L}N(0,\ v(\theta)).$$

而当$\theta=0$时，

$$\lim_{n\to\infty}p_\theta\left\{\sqrt{n}(g^*-\theta)<x\right\}=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}g^*<x\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}\hat{g}<x,|\hat{g}|\geqslant n^{-\frac{1}{4}}\right\}$$

$$+\lim_{n\to\infty}p_\theta\left\{\sqrt{n}a\hat{g}<x,|\hat{g}|<n^{-\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}a\hat{g}<x\right\}-\lim_{n\to\infty}p_\theta\left\{\sqrt{n}a\hat{g}<x,|\hat{g}|\geqslant n^{-\frac{1}{4}}\right\}$$

$$=\lim_{n\to\infty}p_\theta\left\{\sqrt{n}a\hat{g}<x\right\}=N(0,a^2v(\theta)),$$

即当$\theta=0$时，

$$\sqrt{n}(g^*-\theta)\xrightarrow{L}N(0,\ a^2v(\theta)),$$

故(2) 式得证。

由(2) 式取$a\approx 0$，可以使$v^*(\theta)$任意接近于0，从而，g^*所对应的$v^*(\theta)$可低于C—R下界。

特别，设X_1，$X_2\cdots$，X_n为iid样本，$x_1\sim N(\theta,1)$现欲估计θ。θ的极大似然估计为

$$\overline{x}_n=\frac{1}{n}\sum_{i=1}^{n}X_i.$$

$\because\quad \sqrt{n}\left(\overline{x}_n-\theta\right)\xrightarrow{L}N(0,\ 1),$

$\therefore\quad \sqrt{n}\,\overline{x}_n$的渐近方差为 1。但若令

$$\theta_n^* = \begin{cases} 0, & \text{当} \left|\overline{x}_n\right| < n^{-\frac{1}{4}}, \\ \overline{x}_n, & \text{当} \left|\overline{x}_n\right| \geqslant n^{-\frac{1}{4}}. \end{cases}$$

则有 $\sqrt{n}(\theta_n^* - \theta) \xrightarrow{L} N(0, v(\theta))$。

其中 $v(\theta) = \begin{cases} 1, \text{当}\theta \neq 0, \\ 0, \text{当}\theta = 0. \end{cases}$

可见$v(\theta)$可低于c—R下界。

9—14 若$T_n(X)$是θ的一致估计，又$|T_n - \theta_n| \leqslant A_n < \infty$，则$T_n$不是$\theta$的均方一致估计【即$E(T_n - \theta)^2 \nrightarrow 0 (n \to \infty)$】的例子

众所周知，若T_n是θ的均方一致估计（即$E(T_n - \theta)^2 \to 0 (n \to \infty)$，记作$T_n \xrightarrow{2} \theta$），则$T_n$是$\theta$的一致估计（即对$\forall \varepsilon > 0$，$p\{|T_n - \theta| \geqslant \varepsilon\} \to 0 (n \to \infty)$记作$T_n \xrightarrow{p} \theta$）。但其逆不真，即由$T_n \xrightarrow{p} \theta$，却推不出$T_n \xrightarrow{2} \theta$（见〔6〕第一册$p265$）。

很自然，我们要问当$T_n \xrightarrow{p} \theta$时，再加上一点什么条件，就可以推出$T_n \xrightarrow{2} \theta$时，对于这个问题，我们有如下的

定理 设$X_1, X_2, \cdots, X_n$是抽自分布族$\{F(x, \theta), \theta \in \Theta\}$的iid样本，

$$T_n = T_n(X) \triangleq T_n(X_1, X_2, \cdots, X_n),$$

若$T_n \xrightarrow{p} \theta$，且$|T_n - \theta| \leqslant A < \infty$，对$\forall \theta \in \Theta$和$X = (X_1, \cdots, X_n) \in R_n$成立，则$T_n \xrightarrow{2} \theta$。

证明 设$X(X_1 \cdots, X_n)$的联合分布函数为

$$F_n(X_1,\cdots,X_n)\triangleq F_n(X).$$

由于$T_n\xrightarrow{p}\theta$，∴对$\forall\varepsilon>0$，有

$$p\{|T_n-\theta|\geqslant\varepsilon\}\to 0(n\to\infty)\text{对一切}\theta\in\Theta,$$

于是 $$E(T_n-\theta)^2=\int_{R_n}\cdots\int(T_n-\theta)^2dF_n(X)$$

$$=\int\cdots\int_{|T_n-\theta|\geqslant\varepsilon}(T_n-\theta)^2dF_n(X)+\int\cdots\int_{|T_n-\theta|<\varepsilon}(T_n-\theta)^2dF_n(X)$$

$$<A^2\int\cdots\int_{|T_n-\theta|\geqslant\varepsilon}dF_n(X)+\varepsilon^2\int\cdots\int_{|T_n-\theta|<\varepsilon}dF_n(X)$$

$$<A^2p\{|T_n-\theta|\geqslant\varepsilon\}+\varepsilon^2\longrightarrow 0\quad(n\longrightarrow\infty),$$

∴ $T_n\xrightarrow{2}\theta$ 证毕。

进一步的讨论

注 若$T_n\xrightarrow{p}\theta$，但$|T_n-\theta|\leqslant A_n<\infty$时，却推不出$T_n\xrightarrow{2}\theta$

证明 首先我们有如下的结论：存在某个$\varepsilon_0>0$，使得不论N多么大，当$n\geqslant N$时，都有$|T_n-\theta|\geqslant\varepsilon_0$。

事实上，若对$\forall\varepsilon>0$，$\exists N$，当$n\geqslant N$时，有

$$|T_n-\theta|<\varepsilon.$$

取$\varepsilon=1$，并令

$$A=max\{|T_1-\theta|,\cdots,|T_N-\theta|,1\},$$

则对$\forall_n$，有$|T_n-\theta|\leqslant A$，这与假设矛盾。

∴ $$E(T_n-\theta)^2=\int_{R_n}\cdots\int(T_n-\theta)^2dF_n(X)\geqslant\varepsilon^0\not\longrightarrow 0(n\to\infty),$$

故 $$T_n\overset{2}{\not\longrightarrow}0.$$

9—15 极大似然估计不是充分统计量的例子

有关定义和定理

定义 设X具有分布密度族$\{f(x;\ \theta),\theta\in\Theta\}$,其中$\theta=(\theta_1,\cdots,\theta_K)$是一个未知的$K$维参数向量，需待估计．又设$x=(x_1,\cdots,x_n)$是子样$X=(X_1,\cdots,X_n)$的一个观察值．令

$$L(x;\theta)=\prod_{i=1}^{n}f(x_i;\theta)$$（把它看作是θ的函数），则称

$L(x;\theta)$是θ的似然函数。如果选取使下式

$$L(x_1,\cdots,x_n;\hat{\theta})=\sup_{\theta\in\Theta}L(x_1,\cdots,x_n;\theta)$$

成立的$\hat{\theta}(X)=(\hat{\theta}_1(X),\cdots,\hat{\theta}_K(X))$作为$\theta=(\theta_1,\cdots,\theta_K)$的估计，则称$\hat{\theta}(X)$是$\theta$的极大似然估计(离散情形时类似定义)。

关于极大似然估计与充分统计量之间的关系，我们有如下的

定理 设T是分布族$\{f(x;\theta),\ \theta\in\Theta\}$的充分统计量，如果$\theta$的极大似然估计存在,则它是充分统计量$T$的函数(见[6]第二册之二分册$p179$定理1)。

尽管极大似然估计常常是一个充分统计量，但定理并没有断定，极大似然估计本身一定是一个充分统计量．事实上，我们有如下的

例1 设$X=(X_1,X_2,\cdots,X_n)$是抽自服从$U(\theta,\theta+1]$ $(\theta\in R_1)$的母体中的iid样本，则X_1的分布密度为

$$f(x;\theta)=\begin{cases}1, & \theta\leqslant\min\limits_{1\leqslant i\leqslant n}x_i\leqslant\max\limits_{1\leqslant i\leqslant n}X_i\leqslant\theta+1,\\ 0, & \text{其他}.\end{cases}$$

于是，由$R.A.Fisher—J.Neyman$因子分解定理可知$(\min\limits_{1\leqslant i\leqslant n}X_i$,

$\max\limits_{1\leqslant i\leqslant n}X_i$)是$\theta$的联合充分统计量。

又我们知道，对于任意满足

$$\max_{1\leqslant i\leqslant n}X_i-1\leqslant\hat{\theta}\leqslant\min_{1\leqslant i\leqslant n}X_i$$

的$\hat{\theta}$都是θ的极大似然估计，特别

$\min\limits_{1\leqslant i\leqslant n}X_i$是$\theta$的极大似然估计。

但 $\min\limits_{1\leqslant i\leqslant n}X_i$不是$\theta$的充分统计量。

事实上

$$(X_1,\cdots,X_n)\sim\prod_{i=1}^{n}f(x_i;\theta)=\begin{cases}1, & \theta\leqslant\min\limits_{1\leqslant i\leqslant n}X_i\leqslant\max\limits_{1\leqslant i\leqslant n}X_i\leqslant\theta+1,\\ 0, & \text{其他},\end{cases}$$

在$\min\limits_{1\leqslant i\leqslant n}X_i=t$上，有

$$f(x_1,\cdots,x_n\Big|\min_{1\leqslant i\leqslant n}X_i=t)=\begin{cases}\dfrac{\prod\limits_{i=1}^{n}f(x_i;\theta)}{f(t)} & \text{当}\min\limits_{1\leqslant i\leqslant n}X_i=t,\\ 0, & \text{其他}\end{cases}$$

其中$f(t)$为$\min\limits_{1\leqslant i\leqslant n}X_i$的分布密度。

$$f(t)=n[1-t+\theta]^{n-1},\quad \theta\leqslant t\leqslant\theta+1.$$

$$\therefore\quad f(t_1,\cdots,x_n\Big|\min_{1\leqslant i\leqslant n}X_i=t)\begin{cases}\dfrac{1}{n[1-t+\theta]^{n-1}}, & \text{当}\min\limits_{1\leqslant i\leqslant n}X_i=t,\\ 0, & \text{其他}\end{cases}$$

与θ有关，故$\min\limits_{1\leqslant i\leqslant n}X_i$不是$\theta$的充分统计量。

为了举第二个反例，我们需要用到下列预备知识：

设X的分布函数为$F(x)$，密度函数为$f(x)$。

X_1，X_2，X_3相互独立且与X同分布，将它排列为$X_{(1)}\leqslant X_{(2)}\leqslant X_{(3)}$，则$X_{(2)}$有密度函数$6f(y)[1—F(y)]F(y)$。

($X_{(1)}$，$X_{(2)}$)的联合分布密度函数为

$$g(y_1,y_2)=\begin{cases}0, & 当y_1>y_2,\\ 6f(y_1)f(y_2)[1-F(y_2)], & 当y_1\leqslant y_2.\end{cases}$$

例2 设x的密度函数为

$$f(x;\theta)=\frac{1}{2}e^{-|x-\theta|},\quad -\infty<x<+\infty.$$

从中抽取容量为3的iid样本X_1,X_2,X_3，将它排列为$X_{(1)}<X_{(2)}<X_{(3)}$。(X_1,X_2,X_3)的似然函数为

$$\frac{1}{8}e^{-\sum_{i=1}^{3}|x_i-\theta|}$$

θ之极大似然估计为$X_{(2)}$。

下面我们来证明$X_{(2)}$不是θ的充分统计量。

由刚才提到的预备事实，知$x_{(2)}$有密度函数

$$h(x_{(2)}\theta)=6\cdot\frac{1}{2}e^{-|x_{(2)}-\theta|}\int_{-\infty}^{x_{(2)}}\frac{1}{2}e^{-|x-\theta|}dx\int_{x_{(2)}}^{\infty}\frac{1}{2}e^{-|x-\theta|}dx$$

$$=6\cdot\frac{1}{2}e^{-|x_{(2)}-\theta|}c(x_{(2)},\theta)A(x_{(2)},\theta).$$

其中

$$c(x_{(2)},\theta)=\int_{-\infty}^{x_{(2)}}\frac{1}{2}e^{-|x-\theta|}dx=\begin{cases}\dfrac{e^{-\theta}e^{x_{(2)}}}{2}, & 当\theta>x_{(2)}时,\\ \dfrac{1}{2}(2-e^{\theta-x_{(2)}}) & 当\theta\leqslant x_{(2)}时,\end{cases}$$

$$A(x_{(2)},\theta)=\int_{x_{(2)}}^{\infty}\frac{1}{2}e^{-|x-\theta|}dx.$$

$(X_{(1)},X_{(2)})$的联合分布密度为

$$g(X_{(1)},X_{(2)},\theta)=6\cdot\frac{1}{4}e^{-|x_{(1)}-\theta|-|x_{(2)}-\theta|}A(x_{(2)},\theta).$$

故$X_{(1)}|X_{(2)}$的条件分布密度为

$$\frac{g(x_{(1)},\ x_{(2)},\ \theta)}{h(x_{(2)},\theta)}=\frac{1}{2}\frac{e^{-|x_{(1)}-\theta|}}{c(x_{(2)},\theta)}\text{与}\theta\text{有关}$$

$\therefore x_{(2)}$不是θ的充分统计量。

注 极大似然估计是R.A.Fisher在他1912年的一项工作中提出来的。在正态分布这个特殊情况下，这方法可追溯到Gauss在上世纪初关于最小二乘法的工作。Fisher在上述1912年的工作中批评了矩法和最小二乘法，提出了极大似然估计法。

R.A.Fisher很重视极大似然估计，在他的1922年的名著(见本章开头所引)中，尤其是他在1925年发表的《Theory of Statistical Estimation Proc Camb, phil. Soc 22(1925)》一文中对这一估计作了许多研究。

R.A.Fisher更多地是从"极大似然估计能集中样本里多少信息"这个角度去研究这个估计的。那一个时期，他已提出了充分统计量的概念。有一段时期，他相信极大似然估计必是充分统计量，因而在"集中信息"这个角度看是优越的，后来他发现了这一点不对。

尽管如此，Fisher潜心研究所写出的这篇论文仍被认为是参数点估计的奠基性工作。现在一般都认为近代参数点估计的理论始于此文。这是因为Fisher在对极大似然估计的研究中所发展的一些重要概念，诸如一致性，有效性和Fisher信息量等都是首先在这篇论文中提出来的，这对以后的工作有很大的影响。且不论极大似然估计在实用上的极端重要性，就以它在理论上的促进作用来看，也是显著的。可以说，由Fisher开始的工作，直到现在仍是热门的课题。

9—16 极大似然估计不是有效估计的例子

有关定理

下面我们先给出描述有效估计和极大似然估计之间关系的定理：

定理 如果Rao—Cramér不等式定理的条件成立，T是θ的有效估计，则似然方程

$$\frac{2lnL(x_i\theta)}{2\theta}=0$$

具有唯一解T，且T也是θ的极大似然估计。

但定理的逆不一定成立，即并非所有大似然估计（记为MLE：Maximum Likelihood Estimate）一定都是有效估计。请看

例1 设$\mathbf{X}\sim N(\mu,\sigma^2)$其中$\mu$，$\sigma^2$为未知参数，$X=(X_1,\cdots,X_n)$为抽自母体$\mathbf{X}$中的iid样本。

则 $$\hat{\sigma}_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

是σ^2的极大似然估计。但它却不是σ^2的有效估计。

事实上，

$$\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2}\sim x^2(n-1),$$

$$\therefore \quad D\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=2(n-1)\sigma^4,$$

于是 $$D(\hat{\sigma}^2)=\frac{2(n-1)}{n^2}\sigma^4.$$

又∵ $$f(x;\ \sigma^2)=\frac{1}{\sqrt{2\pi}(\sigma^2)^{\frac{1}{2}}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

$$\ln f(x;\sigma^2)=-\frac{1}{2\sigma^2}(x-\mu)^2-\ln\sqrt{2\pi}-\frac{1}{2}\ln\sigma^2$$

∴ $$I(\sigma^2)=E_{\sigma^2}\left(\frac{2\ln f}{2\sigma^2}\right)^2$$

$$=\int_{-\infty}^{\infty}\left[\frac{(x-\mu)^2}{2\sigma^4}-\frac{1}{2\sigma^2}\right]^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\frac{3\sigma^4}{3\sigma^8}-\frac{2\sigma^2}{4\sigma^6}+\frac{1}{4\sigma^4}=\frac{1}{2\sigma^4}.$$

因此，σ^2的无偏估计$\hat{\sigma}^2$的方差下界是

$$\frac{1}{nI(\theta)}=\frac{2\sigma^4}{n},$$

即$\hat{\sigma}^2$的方差 $D(\hat{\sigma}^2)$不等于Rao-Cramér 不等式的下界 $\frac{2\sigma^4}{n}$.

甚至$\hat{\sigma}^2$也不是σ^2的无偏估计. ∴σ^2的极大似然估计

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

不是σ^2的有效估计.

9—17 似然方程的解不是极大似然估计的例子

有关定理

首先，我们叙述正面命题：

定理 若

$$L(x;\ \hat{\theta})=\sup_{\theta\in\Theta}L(x;\theta)=\sup_{\theta\in\Theta}\prod_{i=1}^{n}f(x_i;\ \theta). \qquad (1)$$

其中Θ是开集，且$f(x;\ \theta)$关于θ可微，则满足(1)式的解$\hat{\theta}$(即$\hat{\theta}$是θ的极大似然估计)也一定满足似然方程

$$\left.\frac{\partial \ln L}{\partial \theta}\right|_{\theta=\hat{\theta}}=0.$$

但定理的逆不真，即似然方程的解不一定是极大似然估计。请看

例1 设$X=(X_1,\ \cdots,\ X_n)$是抽自Cauchy分布

$$f(x;\ \theta)=\frac{1}{\pi[1+(x-\theta)^2]} \qquad x\in R_1$$

的iid样本，则

(i) 若$n=1$,则$\hat{\theta}=x_1$。

(ii) 若$n=2$，则似然方程有多重根，但不都是θ的极大似然估计。

事实上，

(i) 若$n=1$，则似然函数

$$L(x_1,\ \theta)=\frac{1}{\pi[1+(x_1-\theta)^2]},$$

于是 $\ln L(x_1;\ \theta)=-\ln\pi-\ln[1+(x_1-\theta)^2]$。

此时，似然方程为

$$\frac{d\ln L}{d\theta}=\frac{2(x_1-\theta)}{1+(x_1-\theta)^2}=0,$$

$\therefore \hat{\theta}=x_1$。但

$$\left.\frac{d^2\ln L}{d\theta^2}\right|_{\theta=\hat{\theta}}<0,$$

故$\hat{\theta}=X_1$是θ的极大似然估计。

(ii) 若$n=2$，则似然方程为

$$L(x_1,x_2,\theta)=\frac{1}{\pi^2[1+(x_1-\theta)^2][1+(x_2-\theta)^2]}. \quad (2)$$

欲使似然函数达到极大，只要(2)的分母达到极小即可．于是，令(2) 的分母为

$$f(\theta)=[1+(x_1-\theta)^2][1+(x_2-\theta^2)], \tag{3}$$

$$\begin{aligned} f'(\theta) &= -2\{(x_1+x_2-2\theta)+(x_1-\theta)(x_2-\theta)[x_1 \\ &\qquad +x_2-2\theta]\} \\ &= -2(x_1+x_2-2\theta)[\theta^2-(x_1+x_2)\theta+x_1x_2+1] \\ &= 0. \end{aligned} \tag{4}$$

解得 $$\theta=\frac{1}{2}(x_1+x_2), \tag{5}$$

$$\theta=\frac{(x_1+x_2)\pm\sqrt{(x_1+x_2)^2-4}}{2}. \tag{6}$$

为判别极小值，考虑$f(\theta)$的二阶导数

$$\begin{aligned} f''(\theta) &= 4[\theta^2-(x_1+x_2)\theta+x_1x_2+1] \\ &\quad -2(x_1+x_2-2\theta)[2\theta-(x_1+x_2)]. \end{aligned} \tag{7}$$

(5)代入(7)得

$$f''\left(\frac{x_1+x_2}{2}\right)=-(x_1-x_2)^2+4. \tag{8}$$

若(8) 式大于等于0，即

$$(x_1-x_2)^2\leqslant 4,$$

即 $$|x_1-x_2|\leqslant 2.$$

则$f(\theta)$达到极小值，而此时(6)为复根，故

$$\hat{\theta}=\frac{1}{2}(x_1+x_2)$$

为θ的极大似然估计，而(6)则不是。若(8)式小于0，则$|x_1-x_2|>2$，

此时 $$f''\left(\frac{x_1+x_2}{2}\right)<0.$$

$f(\theta)$达到极小值．因此

$$\hat{\theta} = \frac{x_1 + x_2}{2}.$$

不是θ的极大似然估计。但在$|x_1 - x_2| > 2$时，(6) 式是两个实数值。将(6) 代入(7) 得

$$f''\left(\frac{(x_1 + x_2) \pm \sqrt{(x_1 - x_2)^2 - 4}}{2}\right)$$

$$= 2[x_1 + x_2 - (x_1 x_2 \pm \sqrt{(x_1 - x_2)^2 - 4})]^2 > 0.$$

故(6) 是$f(\theta)$的极小点，但

$$f\left(\frac{x_1 + x_2 + \sqrt{(x_1 - x_2)^2 - 4}}{2}\right)$$

$$= \frac{1}{16}\left\{4 + \left[(x_1 - x_2) - \sqrt{(x_1 - x_2)^2 - 4}\right]^2\right\}$$

$$\left\{4 + \left[(x_1 - x_2) + \sqrt{(x_1 - x_2)^2 - 4}\right]^2\right\}$$

$$= f\left(\frac{x_1 + x_2 - \sqrt{(x_1 - x_2)^2 - 4}}{2}\right),$$

即(6) 的两个值都使得(2) 成为同一个极小值，它也是$f(\theta)$的最小值。此时(6) 是θ的两个极大似然估计。但不论何种情形，似然方程(4) 有三个根，它们不全是极大似然估计。

例2 设$X = (X_1, \cdots, X_n)$是抽自均匀分布

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{当}\alpha < x < \beta, \\ 0, & \text{其他}. \end{cases}$$

的iid样本。其中α和β是使$\alpha < \beta$的两个数。似然函数为

$$L(\alpha; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{当}\alpha < x < \beta, \\ 0, & \text{其他}. \end{cases}$$

如果我们从似然方程

$$\frac{\partial L(x;\alpha,\ \beta)}{\partial \alpha}=0,$$

$$\frac{\partial L(x;\ \alpha,\ \beta)}{\partial \beta}=0$$

来求解α和β，则我们发现α和β至少有一个必是无穷大，这是无意义的结果．造成困难的原因在于似然函数在极大值处没有零斜率。

但我们不难从其他途径找到α、β的极大似然估计分别为

$$\hat{\alpha}=\min(X_1,\ \cdots,\ X_n),\quad \hat{\beta}=\max(X_1,\ \cdots,\ X_n).$$

进一步的讨论

注 设$X=(X_1,\ X_2,\ \cdots,\ X_n)$是取自分布族$\{f(x;\ \theta),\ \theta\in\Theta\}$的iid样本，$T(X)$是$\theta$的有效估计，$f(x_i;\ \theta)$和$T$满足 Rao-Cramér 不等式定理的条件，则似然方程$\frac{\partial \ln L(x;\ \theta)}{\partial \theta}=0$具有唯一的解$T(X)$，且$T(X)$必也是$\theta$的极大似然估计（见〔6〕第二册之二分册$P180$定理2）。

(*)9—18 极大似然估计不是一致估计的例子

关于极大似然估计的相合性问题，引起了许多统计学者的兴趣．直到现在都不能说这个问题已经彻底解决了．1946 年 Cramér在一些条件下证明了似然方程有一个根是参数θ的弱相合估计(见例9—13)．由于似然方程的根不一定就是极大似然估计(见例9—17)，所以这个结果还没有解决极大似然估计的相合性问题．直到1949年wald才首次证明了极大似然估计的强相合性，但所要求的条件很复杂．在这里我们提供一个反例，说明极大似然估计确实可以不相合。

有关定理

在给出反例之前，我们先来叙述上面提到的wald定理：

定理 设随机变量X的样本空间和分布族分别为$(\mathscr{X}, \beta_*)$和$\{P_\theta, \ \theta \in \Theta\}$，$\Theta = R_k$，假定

$$\{P_\theta, \ \theta \in \Theta\} \ll \mu$$

（即P_θ关于μ绝对连续），μ为β_*上的σ有限测度，记

$$f(x; \ \theta) = \frac{dP_\theta(X)}{d\mu}$$

以及

$$f(x; \ \theta, \ \zeta) = \sup_{\|\theta' - \theta\| < \zeta} f(x; \ \theta') \qquad (0 < \zeta \leqslant \infty),$$

$$\varphi(x, \ r) = \sup_{\|\theta\| > r} f(x; \ \theta) \qquad (0 \leqslant r < \infty),$$

$$f^*(x, \ \theta, \ \zeta) = \max\{f(x, \ \theta, \ \zeta), \ 1\},$$

$$\varphi^*(x, \ r) = \max\{\varphi(x, \ r), \ 1\}.$$

又假定

(i) 对每个$\theta \in \Theta$和$\zeta > 0$，$f(x, \ \theta, \ \zeta)$为β_*可测；

(ii) 对每个$\theta \in \Theta$及$\theta_0 \in \Theta$，存在$\zeta_{\theta, \theta_0} > 0$使当$0 < \zeta < \zeta_{\theta, \theta_0}$时，有

$$\int_* [\ln f^*(x, \ \theta, \ \zeta)] f(x; \ \theta_0) d\mu(x) < \infty.$$

又当r充分大（与θ_0有关）时

$$\int_* [\ln \varphi^*(x, \ r)] f(x; \ \theta_0) d\mu(x) < \infty.$$

(iii) 对任何$\theta \in \Theta$有

$$\int_* |\ln f(x, \ \theta)| f(x, \ \theta) d\mu(x) < \infty.$$

(iv) 存在一集$A \in \beta_*$，使对任何θ有

$$\int_A f(x; \ \theta) d\mu(x) = 0,$$

且 $\lim\limits_{\|\theta'\|\to\infty} f(x;\ \theta')=0$ 当$x\in A$

(v) 当$\theta_1\neq\theta_2$时，

$\mu(\{x;\ f(x;\ \theta_1)\neq f(x;\ \theta_2)\})>0$，

(vi) 对任何$\theta\in\Theta$，存在一集$B_\theta\in\beta_*$，使

$\int_{B_\theta} f(x,\ \theta_0)d\mu=0$。

对任何$\theta_\theta\in\Theta$，且当$x\in B_\theta$时，有

$f(x;\ \theta')\longrightarrow f(x;\ \theta)$，

对任何$\theta'\longrightarrow\theta$。又设$X_1$，…，$X_n$为抽自母体$\mathbf{X}$的iid样本，且对任何$n$，基于样本$X_1$，…，$X_n$的$\theta$的极大似然估计$\theta_n^*(X_1$，…，$X_n)$存在，则$\theta_n^*$是$\theta$的强一致估计（当然更是$\theta$的一致估计）（见〔7〕p206定理2.6.5）

然而，确实存在着极大似然估计不是一致估计的例子。下面的例子是Basu给出的。

例1 设$X=(X_1$，…，$X_n)$为抽自母体X的iid样本，X的分布为

$$P(X=x)=\begin{cases}\theta, & x=1,\ \theta\text{为有理数},\\ 1-\theta, & x=1,\ \theta\text{为无理数}.\end{cases}$$

$$P(X=x)=\begin{cases}1-\theta, & x=0,\ \theta\text{为有理数},\\ \theta, & x=0,\ \theta\text{为无理数}.\end{cases}$$

统一写成

$$P(X=x)=f(x;\ \theta)=\begin{cases}\theta^x(1-\theta)^{1-x}, & \theta\text{ 为有理数},\\ (1-\theta)^x\theta^{1-x}, & \theta\text{ 为无理数}.\end{cases}$$

此处$0\leqslant\theta\leqslant1$。以$T_n$为$X_1+X_2+\cdots+X_n$，则$(X_1$，…，$X_n)$的似然函数为

$$L(x_1, \cdots, x_n; \theta) = \begin{cases} \theta^{\sum\limits_{i=1}^{n} X_i}(1-\theta)^{n-\sum\limits_{i=1}^{n} X_i}, \\ \quad \theta\text{为有理数}, \\ (1-\theta)^{\sum\limits_{i=1}^{n} X_i}\theta^{n-\sum\limits_{i=1}^{n} X_i}, \\ \quad \theta\text{为无理数} \end{cases}$$

$$= \begin{cases} \theta^{T_n}(1-\theta)^{n-T_n}, & \theta\text{为有理数}, \\ (1-\theta)^{T_n}\theta^{n-T_n}, & \theta\text{为无理数}. \end{cases}$$

$$\frac{\partial L}{\partial \theta} = T_n\theta^{T_n-1}(1-\theta)^{n-T_n} - (n-T_n)(1-\theta)^{n-T_n-1}\theta^{T_n}$$

$$= 0,$$

得 $$\theta^{T_n-1}(1-\theta)^{n-T_n-1}[(1-\theta)T_n - (n-T_n)\theta] = 0,$$

∴ $$\theta_n^* = \frac{T_n}{n}.$$

又 $$\frac{\partial^2 L}{\partial \theta^2} = T_n(T_n-1)\theta^{T_n-2}(1-\theta)^{n-T_n}$$

$$-(n-T_n)(1-\theta)^{n-T_n-1}T_n\theta^{T_n-1}$$

$$+(n-T_n)(n-T_n-1)(1-\theta)^{n-T_n-2}\theta^{T_n}$$

$$-T_n\theta^{T_n-1}(n-T_n)(1-\theta)^{n-T_n-1}$$

$$= \theta^{T_n-2}(1-\theta)^{n-T_n-2}[T_n(T_n-1)(1-\theta)^2$$

$$-(n-T_n)T_n(1-\theta)\theta + (n-T_n)(n-T_n-1)\theta^2$$

$$-T_n\theta(1-\theta)(n-T_n)]$$

$$\left.\frac{\partial^2 L}{\partial \theta^2}\right|_{\theta=\frac{T_n}{n}} = \left(\frac{T_n}{n}\right)^{T_n-2}\left(1-\frac{T_n}{n}\right)^{n-T_n-2}$$

$$\times\left[T_n(T_n-1)\left(1-\frac{T_n}{n}\right)^2\right.$$

$$-(n-T_n)T_n\left(1-\frac{T_n}{n}\right)\frac{T_n}{n}$$

$$+(n-T_n)(n-T_n-1)\left(\frac{T_n}{n}\right)^2$$

$$-T_n\frac{T_n}{n}\left(1-\frac{T_n}{n}\right)(n-T_n)\Big]$$

$$=\left(\frac{T_n}{n}\right)^{T_n-2}\left(1-\frac{T_n}{n}\right)^{n-T_n-2}$$

$$\times\Big[\frac{T_n(T_n-1)}{n^2}(n-T_n)^2$$

$$+(n-T_n)(n-T_n-1)\frac{T_n^2}{n^2}$$

$$-2\frac{T_n^2}{n^2}(n-T_n)^2\Big]$$

$$<\left(\frac{T_n}{n}\right)^{T_n-2}\left(1-\frac{T_n}{n}\right)^{n-T_n-2}\Big[\frac{T_n^2}{n^2}(n-T_n)^2$$

$$+\frac{T_n^2(n-T_n)^2}{n^2}-2\frac{T_n^2}{n^2}(n-T_n)^2\Big]=0,$$

∴ 当θ是有理数时

$$\sup_{0\leqslant\theta\leqslant1}\theta^{T_n}(1-\theta)^{n-T_n}$$

在有理点$\theta_n^*=\frac{T_n}{n}$处达到极大值，其极大值为

$$\left(\frac{T_n}{n}\right)^{T_n}\left(1-\frac{T_n}{n}\right)^{n-T_n}.$$

同理，当θ是无理数时，$(1-\theta)^{T_n}\theta^{n-T_n}$在$\theta<1-\frac{T_n}{n}$时上升，在$\theta>1-\frac{T_n}{n}$时下降。

故 $\sup\limits_{0\leq\theta\leq 1}(1-\theta)^{T_n}\theta^{n-T_n}$也在有理点$\theta_n^*=1-\dfrac{T_n}{n}$处达到极大值。
故θ为无理数时，其极大值仍为

$$\left(\frac{T_n}{n}\right)^{T_n}\left(1-\frac{T_n}{n}\right)^{n-T_n}.$$

故$\theta_n^*=\dfrac{T_n}{n}$为$\theta$的极大似然估计。

又当θ为有理数时

$$E_\theta X_1=\theta.$$

θ为无理数时

$$E_\theta X_1=1-\theta,$$

故由强大数律

$$\frac{T_n}{n}=\frac{1}{n}\sum_{i=1}^{n}X_i\longrightarrow E_\theta X_1=\theta,\text{ 当}\theta\text{为有理数时，}$$

$$\frac{T_n}{n}=\frac{1}{n}\sum_{i=1}^{n}X_i\longrightarrow E_\theta X_1=1-\theta,\text{ 当}\theta\text{为无理数时。}$$

即
$$\lim_{n\to\infty}\theta_n^*=\begin{cases}\theta, & \text{当}\theta\text{为有理数，}\\ 1-\theta, & \text{当}\theta\text{为无理数.}\end{cases}$$

这说明极大似然估计θ_n^*不是θ的一致估计。

第十章　与假设检验有关的反例

假设检验是数理统计中另一类重要的推断问题．本章是在假定读者已经熟知假设检验理论中诸如势函数、检验函数、检验水平、单边假设、双边假设、一致最优势检验及一致最优势无偏检验等等概念的前提下进行讨论的．

10—1　不是单参数指数族，但具有单调似然比的例子

有关定义和定理

定义　对于分布密度族$\{f(x;\ \theta)\}$，其中θ是实参数．如果存在实值函数$T(x)$，使对任意$\theta_1<\theta_2$，满足

(1) $P\{f(x;\ \theta_1)\neq f(x;\ \theta_2)\}>0$．

(2) 似然比

$$\lambda(x)=\frac{f(x;\ \theta_2)}{f(x;\ \theta_1)}$$

是关于$T(x)$的非降函数（或非增函数），则称分布密度族$\{f(x;\ \theta)\}$关于$T(x)$具有单调似然比．

如果X是离散型的随机变量，对其概率分布族$\{P(x;\ \theta)\}$具有定义中所叙述的性质，也同样称$\{P(x;\ \theta)\}$关于$T(x)$具有单

调似然比。

关于单参数指数族，我们有如下的

定理 设子样$X=(X_1, \cdots X_n)$服从单参数指数族分布，其联合分布密度为

$$f(x; \theta) = a(\theta)\exp\{Q(\theta)T(x)\}h(x). \qquad (1)$$

其中θ是一个实参数，$Q(\theta)$是关于θ的严格单调增加函数，则分布族$\{f(x; \theta)\}$关于$T(x)$具有单调似然比（见〔6〕第二册之二分册P218）。

但是，并非单参数指数族才具有单调似然比。我们有如下的

例1 设$X_1, X_2, \cdots, X_n$ iid, $X_1 \sim U[0, \theta]$ $\theta>0$. 则$(X_1, \cdots, X_n)$的联合分布密度为

$$f(x; \theta)=\begin{cases}\dfrac{1}{\theta^n}, 当 0 \leqslant \max(X_1, \cdots, X_n) \leqslant \theta, \\ 0, 其他.\end{cases}$$

设$\theta_1<\theta_2$，考虑似然比

$$\frac{f(x; \theta_2)}{f(x; \theta_1)}=\frac{\left(\dfrac{1}{\theta_2^n}\right) I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_2]}{\left(\dfrac{1}{\theta_1^n}\right) I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_1]}$$

$$=\left(\frac{\theta_1}{\theta_2}\right)^n \frac{I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_2]}{I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_1]}.$$

令 $$k(X)=\frac{I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_2]}{I[\max\limits_{1 \leqslant i \leqslant n} X_i \leqslant \theta_1]}$$

$$=\begin{cases}1, 当 \max\limits_{1 \leqslant i \leqslant n} X_i \in [0, \theta_1], \\ \infty, 当 \max\limits_{1 \leqslant i \leqslant n} X_i \in [\theta_1, \theta_2].\end{cases}$$

定义$R(X)=\infty$，如果$\max\limits_{1\leqslant i\leqslant n} X_i>\theta_2$

由此可知，$\dfrac{f(x;\ \theta_2)}{f(x;\ \theta_1)}$是$\max\limits_{1\leqslant i\leqslant n} X_i$的非降函数，所以，在$[0,\ \theta]$上的均匀分布族（它不属于单参数指数族）关于$\max\limits_{1\leqslant i\leqslant n} X_i$具有单调似然比。

进一步的讨论

注 并非所有的分布族都具有单调似然比。

下面是一个不存在单调似然比的分布族的例子。

例2 设

$$x\sim\frac{1}{\pi[1+(x-\theta)^2]}\qquad\text{（哥西分布）}$$

设$\theta_2>\theta_1$，则

$$\frac{f(x;\ \theta_2)}{f(x;\ \theta_1)}=\frac{1+(x-\theta_1)^2}{1+(x-\theta_2)^2}\longrightarrow 1$$

（当$x\longrightarrow\pm\infty$时）

因此哥西（Cauchy）分布族$\left\{\dfrac{1}{\pi[1+(x-\theta)^2]}\right\}$不存在单调似然比。

(*)10—2 单边假设检验不存在一致最优势检验（记为UMP检验）的例子

我们知道，在有些情况下，关于

$$H_0:\ \theta=\theta_0\longleftrightarrow H_1:\ \theta>\theta_0（或\theta<\theta_0）.$$

这类单边假设检验问题，存在着UMP检验．最典型的例子有：

设母体服从正态分布$N(\mu, 1)$，其中μ是未知参数。欲检验假设

$$H_0: \mu=0 \longleftrightarrow H_1: \mu>0, \tag{1}$$

则不难知道检验函数

$$\phi(x)=\begin{cases}1, & 当\bar{x}\geqslant A时,\\ 0, & 当\bar{x}<A时.\end{cases}$$

其中A满足

$$\int_A^\infty \frac{\sqrt{n}}{\sqrt{2\pi}}e^{-\frac{nt^2}{2}}dt=\alpha \qquad (0<\alpha<1)$$

是上述单边假设检验问题(1)的UMP检验。

但在一般情况下，对于单边假设检验问题，由于最优势检验（MP检验）依赖于H_1假设中θ的具体值，因此UMP检验可以不存在。那么，什么样的单边假设检验问题，其UMP检验就不存在呢？下面我们给出一个例子。

例1 Cauchy测量模型。

设$X_1, \cdots, X_n$是对某个量Δ的n次独立测定。我们有$X_i=\Delta+\varepsilon_i \quad i=1, 2, \cdots, n$

其中随机误差$\varepsilon_i=(i=1,2,\cdots, n)$是$n$个相互独立的随机变量，并且假定$\varepsilon_i$（$i=1, 2, \cdots, n$）都服从Cauchy分布

$$f(x)=\frac{1}{\pi(1+x^2)}, \quad -\infty<x<+\infty.$$

我们考虑如下的单边假设检验问题：

$$H_0: \Delta=0 \longleftrightarrow H_1: \Delta>0. \tag{2}$$

下面我们来说明，对这个单边假设检验问题(2)，的确不存在UMP检验。因为x_i的分布密度为

$$p(x_i\ \Delta)=\frac{1}{\pi[1+(x_i-\Delta)^2]},$$

所以X_1，X_2，…，X_n的联合分布密度为

$$p(x_1, \cdots, x_n; \Delta) \triangleq p(x; \Delta) = \frac{1}{\pi^n}\prod_{i=1}^{n} \times \frac{1}{[1+(x_i-\Delta)^2]}.$$

由此，可得到似然比为

$$L(x, 0, \Delta) = \frac{\prod_{i=1}^{n} p(x_i; \Delta)}{\prod_{i=1}^{n} p(x_i; 0)} = \prod_{i=1}^{n} \frac{1+x_i^2}{1+(x_i-\Delta)^2}.$$

如果问题(2)有UMP检验，则必对任何$\Delta_2 > \Delta_1 > 0$

$L(x, 0, \Delta_2)$和$L(x, 0, \Delta_1)$

将生成同样一族临界域．意即对任何$k_1 \geqslant 0$，必有$k_2 \geqslant 0$；对任何$k_2 \geqslant 0$，必有$k_1 \geqslant 0$，使得

(i) $\{x: L(x, 0, \Delta_1) \geqslant k_1\} = \{x: L(x, 0, \Delta_2) \geqslant k_2\}$.

下面我们来看这有没有可能呢？

首先来看$n=1$（即只有一个样本）的情形，

这时 $L(x, 0, \Delta) = \dfrac{p(x; \Delta)}{p(x; 0)} = \dfrac{1+x^2}{1+(x-\Delta)^2}$.

欲使条件(i)成立，必须对任何$x_1 \neq x_2$，有关系

(ii) $L(x_1, 0, \Delta_1) \leqslant L(x_2, 0, \Delta_1) \Longrightarrow L(x_1, 0, \Delta_2) \leqslant L(x_2, 0, \Delta_2)$（条件(ii)表明$L(x, 0, \Delta_1)$增加，必有$L(x, 0, \Delta_2)$增加）这是因为，若

$L(x_1, 0, \Delta_2) > L(x_2, 0, \Delta_2)$,

那么，取$k_2 = L(x_1, 0, \Delta_2)$，则对任何$k_1$，条件(i)不成立．但由于

$$L'_x(x, 0, \Delta)=\frac{-2\Delta(x^2-\Delta\cdot x-1)}{[1+(x-\Delta)^2]^2}$$

对于$\Delta_2>\Delta_1>0$，在区间$\left(\frac{\Delta_1}{2}, \frac{\Delta_2}{2}\right)$上$L(x,0, \Delta_1)$是减函数，而$L(x, 0, \Delta_2)$是增函数，因此，$L(x, 0, \Delta)$不满足条件(ii)，因而条件(i)也不满足。这样，我们证明了在$n=1$时，单边假设检验问题(2)不存在UMP检验。

一般情形，我们只要说明$n=3$就够了。这相当于(ii)的条件是：对任意固定的两个坐标，譬如$x_1=x_1^0$，$x_3=x_3^0$，

$$L(x_1^0, x_2, x_3^0, 0, \Delta_1)\text{ 和 }L(x_1^0, x_2, x_3^0, 0, \Delta_2)$$

关于x_2满足条件(ii)。由此，并注意到$L(x, 0, \Delta)$的形式，再利用$n=1$的结果，即可知假设检验问题(2)不存在UMP检验。

(*)10—3 双边假设检验不存在UMP检验的例子

我们知道，若$z=(z_1, \cdots, z_n)$服从单参数指数族分布，其分布密度为：

$$f(x; \theta)=a(\theta)exp\{Q(\theta)T(x)\}h(x).$$

其中θ是一个实值参数，$Q(\theta)$是θ的严格单调增加函数，则对于双边假设检验

H_0: $\theta\leqslant\theta_1$或$Q\geqslant\theta_2$ $(\theta_2>\theta_1)\longleftrightarrow H_1$: $\theta_1<\theta<\theta_2$问题，

对任意水平$\alpha(0<\alpha<1)$存在UMP检验。

但是，并非所有的双边假设检验问题都存在UMP检验。下面是两个例子。

例1 设$x\sim N(\mu, \sigma^2)$，其中σ^2已知，欲检验

$$H_0: \mu=\mu_0\longleftrightarrow H_1: \mu\neq\mu_0, \tag{1}$$

则对于这个双边假设检验问题(1)，对任何α都不存在水平为α的UMP检验．这是因为，对单边假设检验问题

$$H_0: \mu=\mu_0 \longleftrightarrow H_1: \mu>\mu_0 \qquad (2)$$

而言，检验

$$\phi_1(x)=\begin{cases}1, & 当\bar{x}\geqslant\mu_0+A\dfrac{\sigma}{\sqrt{n}}\\ 0, & 其他\end{cases}$$

（其中A由

$$\int_A^\infty \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} e^{-\frac{n(t-\mu_0)^2}{2\sigma^2}}dt=\alpha$$

所确定）是其水平为α的UMP检验，其势函数是

$$\beta_{\Phi_1}(\mu)=E_\mu\Phi_1=p_\mu\left\{\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}\geqslant A\right\}$$

$$=p\left\{\frac{\sqrt{n}\bar{X}}{\sigma}\geqslant A+\frac{\sqrt{n}\mu_0}{\sigma}\right\}=p\left\{\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\right.$$

$$\left.\geqslant A+\frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right\}$$

$$=1-p\left\{\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}<A+\frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right\}$$

$$=1-\Phi\left(A+\frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right)=\Phi\left(\frac{\sqrt{n}(\mu-\mu_0)}{\sigma}-A\right),$$

$$\mu\geqslant\mu_0.$$

而对单边假设检验问题

$$H_0: \mu=\mu_0 \longleftrightarrow H_1: \mu<\mu_0 \qquad (3)$$

而言，检验

$$\phi_2(x)=\begin{cases}1, & 当\bar{X}\leqslant\mu_0-A\dfrac{\sigma}{\sqrt{n}},\\ 0, & 其他\end{cases}$$

（其中A仍由

$$\int_A^{\infty}\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{n(t-\mu_0)^2}{2\sigma^2}}dt=\alpha$$

所确定）是其水平为α的UMP检验，其势函数为

$$\beta_{\phi_2}(\mu)=E_\mu\phi_2=\Phi\left(\frac{\sqrt{n}(\mu_0-\mu)}{\sigma}-A\right),\quad \mu\leqslant\mu_0.$$

势函数$\beta_{\phi_1}(\mu)$和$\beta_{\phi_2}(\mu)$的图形如下图所示.

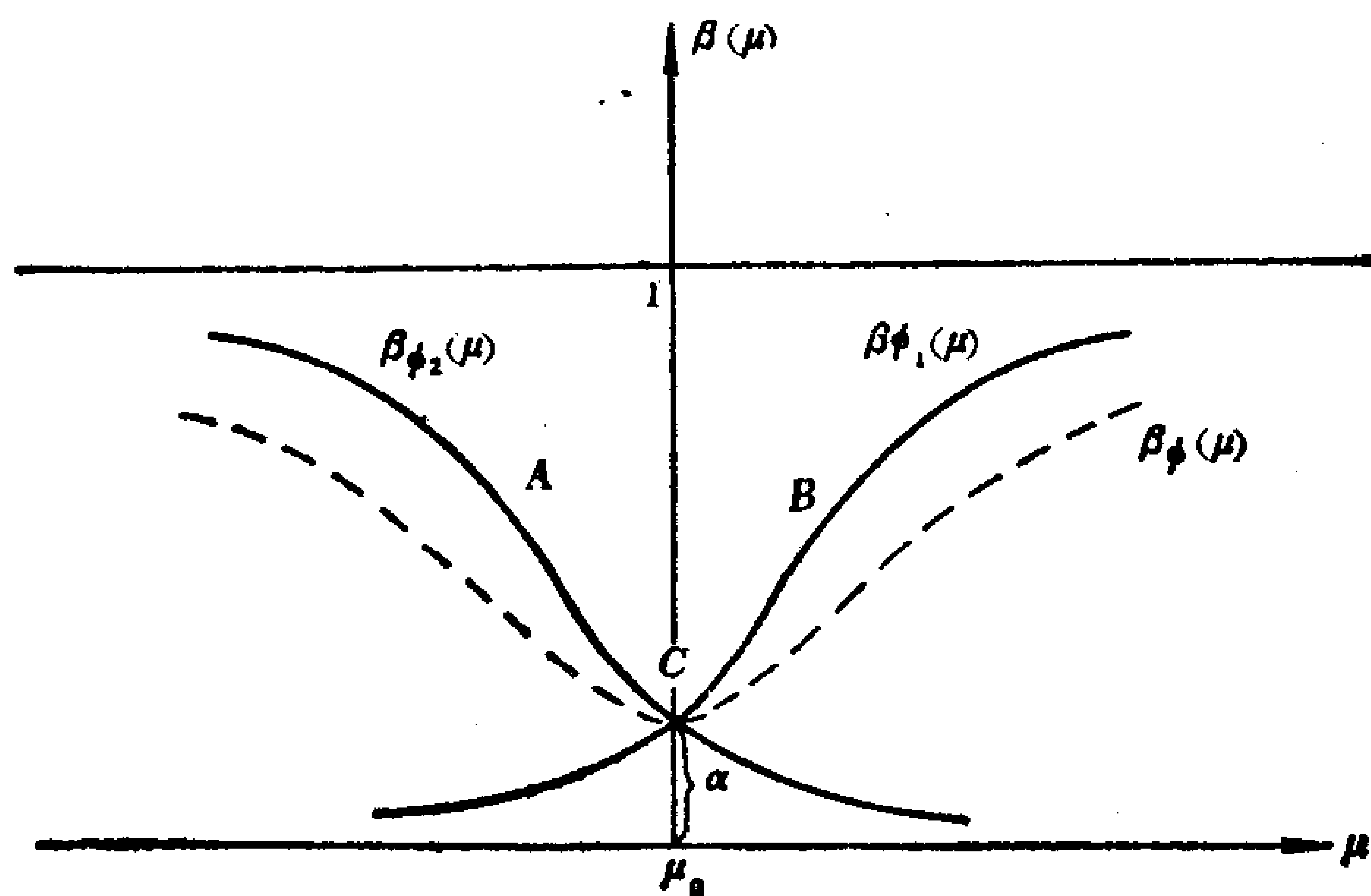


图10—3(1)

现在，我们来证明双边假设检验问题（1）不存在水平α的UMP检验.

反证法.如果双边假设检验问题(1)存在水平为α的UMP检验$\phi(X)$，则其势函数

$$\beta_\phi(\mu)=E_\mu(\phi)$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\int_{R_n}\cdots\int\phi(x_1\cdots x_n)e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

必与图中的曲线$A—B—C$相重合。但是曲线$A—C—B$在C点无导数，而$\beta_\phi(\mu)$关于μ是处处可导的，这得到矛盾。所以，对双边假设检验问题(1) 不存在检验函数以$A—C—B$为其势函数。

例2 对$x \sim N(0,\sigma^2)$，$X_1\cdots$，X_n是从x中抽取的iid样本。因为$X=(X_1,\cdots,X_n)$的联合分布密度为

$$f(X;\sigma^2)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2\right\}$$

$$=a(\sigma^2)\exp\{T(x)Q(\sigma^2)\}.$$

它是单参数指数族分布，其中

$$T(x)=\sum_{i=1}^{n}X_i^2,\qquad Q(\sigma^2)=-\frac{1}{2\sigma^2}.$$

显然$Q(\sigma^2)$ 是σ^2的单调上升函数，又$f(x;\sigma^2)$关于$T(x)$具有单调似然比，所以，对于单边假设检验问题

$$H_0:\ \sigma\leqslant\sigma_0\longleftrightarrow H_1:\ \sigma>\sigma_0 \tag{4}$$

和

$$H_0:\ \sigma\geqslant\sigma_0\longleftrightarrow H_1:\ \sigma<\sigma_0 \tag{5}$$

都存在水平为α的UMP检验。

但是，对于双边假设检验问题

$$H_0:\ \sigma=\sigma_0\longleftrightarrow H_1:\ \sigma\neq\sigma_0 \tag{6}$$

却不存在水平为α的UMP检验。

这是因为，对于单边假设检验问题

$$H_0:\ \sigma=\sigma_0\longleftrightarrow H_1:\ \sigma>\sigma_0. \tag{7}$$

检验

$$\phi_1(x)=\begin{cases}1 & \text{当}\sum_{i=1}^{n}x_i^2>C_1\\ 0 & \text{其他}\\ (\text{其中}C_1=\sigma_0^2x_{n,\alpha}^2) & \end{cases}$$

是其水平为α的UMP检验，其势函数为

$$\beta_{\phi_1}(\sigma^2)=p_{\sigma_1}\left(\sum_{i=1}^{n}x_i^2>C_1\right) \qquad \sigma>\sigma_0$$

而对于单边假设检验问题

$$H_0:\ \sigma=\sigma_0\longleftrightarrow H_1:\ \sigma<\sigma_0 \tag{8}$$

检验

$$\phi_2(x)=\begin{cases}1 & \text{当}\ \sum_{i=1}^{n}x_i^2<C_2\\ 0 & \text{其他}\end{cases}$$

（其中 $C_2=\sigma_0^2\chi^2_{n,1-\alpha}$）

是其水平为α的UMP检验，其势函数为

$$\beta_{\phi_2}(\sigma^2)=\rho_{\sigma_2}\left(\sum_{i=1}^{n}x_i^2<C_2\right), \qquad \sigma<\sigma_0$$

势函数$\beta_{\phi_1}(\sigma^2)$与$\beta_{\phi_2}(\sigma^2)$的图形如下图所示：

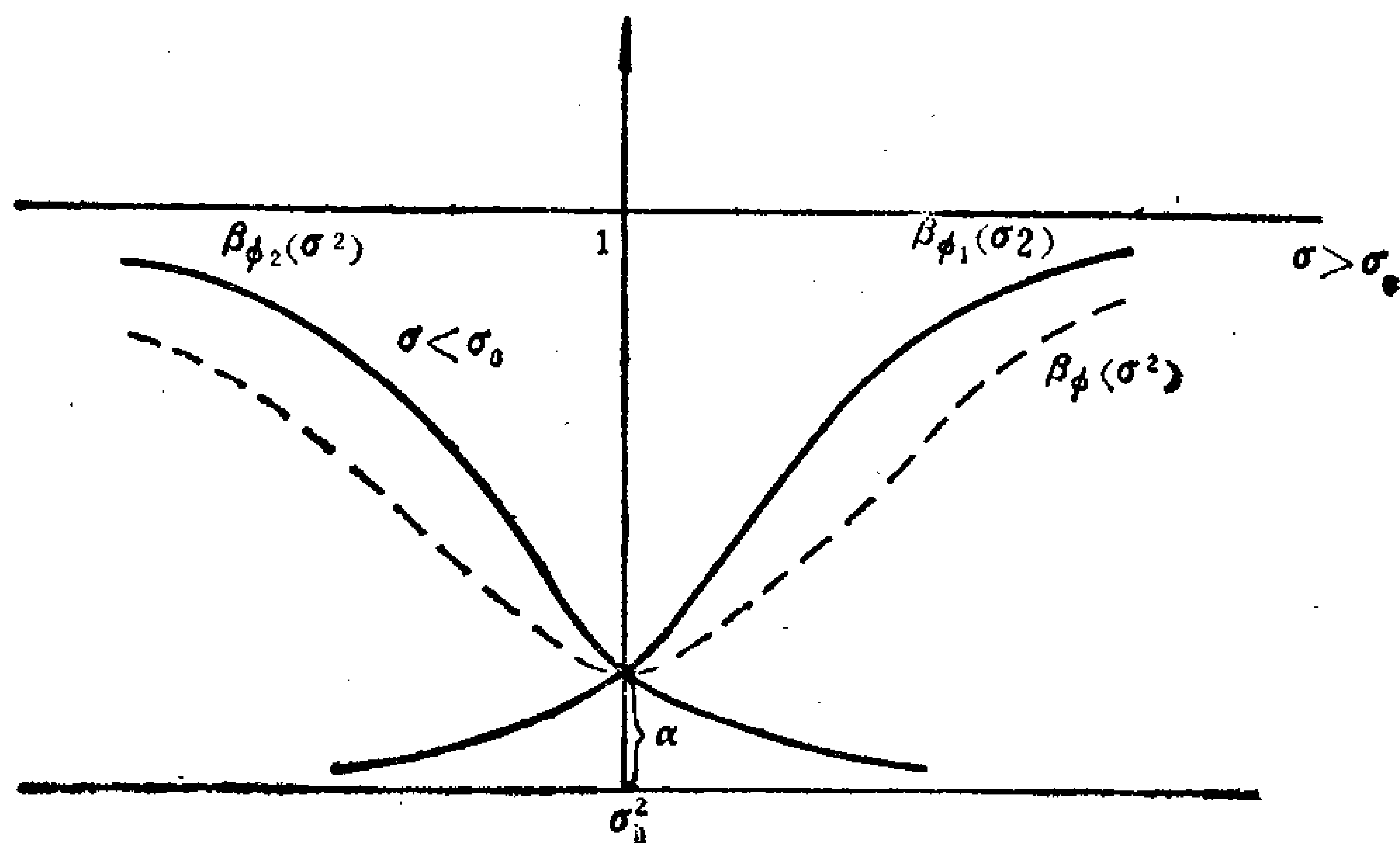


图10—3(2)

显然，类似例1的说明可知，对双边假设检验问题(6)不存在水平为α的UMP检验。

(*)10--4 检验函数ϕ对公共边界相似，但ϕ没有Neyman结构的例子

有关定义和定理

定义1 若ϕ为假设检验问题

$$H_0: \theta\in\Theta_0 \longleftrightarrow H_1: \theta\in\Theta_1=\Theta-\Theta_0$$

的一个检验，其势函数$\beta_\phi(\theta)$之值在ω上保持不变，ω为Θ之一子集，则称检验ϕ对于集ω相似，或者说，ϕ关于分布族$\{p_\theta,\theta\in\omega\}$相似。

定义2 设$t(x)$为关于分布族$\{p_\theta,\theta\in\omega\}$的一个充分统计量(这意味着固定$A$时，对一切$\theta\in\omega$，可选出一个公共的$p_\theta(A|t)$，因而，给定检验$\phi$，对一切$\theta\in\omega$，可选出一个公共的$E_\theta(\phi(X)|t)$，这记为$E(\phi(X)|t)$)。如果存在$\alpha$，$0\leqslant\alpha\leqslant1$，使

$E(\phi(X)|t)=\alpha$，$(a、e、p_\theta)$对任何$\theta\in\omega$，则称ϕ对(t,ω)有Neyman结构。

关于ϕ对(t,ω)有Neyman结构，与ϕ对ω相似之间的关系，我们有以下结论：

定理1 若ϕ对(t,ω)有Neyman结构，则ϕ对ω相似（见[7] P256—257)。

但定理的逆不真。即若ϕ对ω相似，但它不一定对(t,ω)有Neyman结构。请看

例1 设$X\sim N(\theta,1)$　　$\Theta=(-\infty,\infty)$

$\omega=\{0\}$　样本大小为1，

$t(x)=x$　为θ的充分统计量，

任给α，$0<\alpha<1$。

取　　H_0：$\theta=0 \longleftrightarrow H_1$：$\theta \neq 0$

之检验函数为

$$\phi(x)=\begin{cases}1, & 当|x|>u_{\frac{\alpha}{2}}, \\ 0, & 当|x| \leqslant u_{\frac{\alpha}{2}}.\end{cases}$$

则 ϕ为水平α的相似检验．但

$$E_\theta[\phi(X)|x]=\varphi(x) \neq \alpha \qquad 对\theta \in \omega=\{0\},$$

故ϕ对 (t,ω)无Neyman结构.

进一步的讨论

注　但是有如下的结论：

定理2　设$t(x)$为关于分布族$\{p_\theta,\theta \in \omega\}$的一个充分且有界完全统计量，则任一检验$\phi$对$(t,\omega)$有Neyman 结构的充要条件为它对$\omega$相似（见〔7〕P257定理3.3.1）.

10—5　当正则条件不成立时，Wilks 定理中关于似然比极限分布的结论可以不成立的例子

有关概念和定理

设随机变量X的分布族为

$$\{f(x;\theta)d\mu(x),\theta=(\theta_1\cdots\theta_k) \in \Theta\}. \tag{1}$$

Θ为R_k中的一个有内点的集合.

假设分布族(1) 满足正则性条件：

(i) $\displaystyle\int_* \frac{\partial f(x;\theta)}{\partial\theta_i}d\mu(x)=0, i=1,2,\cdots,k.$

(ii) $I(\theta)=(I_{ij}(\theta))_{k\times k}>0$，对任何$\theta \in \Theta$.

其中　$$I_{ij}(\theta)=\int_*\left(\frac{\partial \log f(x;\theta)}{\partial\theta_i}\right)\left(\frac{\partial \log f(x;\theta)}{\partial\theta_j}\right)f(x;\theta)d\mu(x)$$

(iii) 存在$M(x)$，使$\int M(x)f(x;\theta)\,d\mu(x)<\infty$，对任何$\theta\in\Theta$。且

$$\left|\frac{\partial^3 \log f(x;\theta)}{\partial\theta_a\partial\theta_b\partial\theta_c}\right|\leqslant H(x),a,b,c=1,2,\cdots,k.$$

(iv) 不同的θ值相应于X的不同分布，考虑假设

$$H_0:\ \theta\in\Theta_0\longleftrightarrow H_1:\ \theta\in\Theta-\Theta_0$$

关于H_0假设的集合Θ_0，我们假定它是r维的，即存在R_r中的一个有内点的集合A，以及定义在A上的k个函数g_i：

$$\theta_i=g_i(\varphi_1,\cdots,\varphi_r),i=1,\cdots,k.$$

它们建立了A与Θ_0的一一对应，要求每个g_i在A的内点处有直到三阶为止的偏导数，这样，

分布族 $\{f(x;\theta)d\mu(x),\ \theta\in\Theta_0\}$

与分布族

$$\begin{aligned}\{\tilde{f}(x,\varphi)d\mu=\tilde{f}(x,\varphi_1,\cdots,\varphi_r)d\mu(x)\\ =f(x,g_1(\varphi_1,\cdots,\varphi_r),\cdots,g_k(\varphi_1,\cdots,\varphi_r))\\ d\mu(x),\varphi\in A\}\end{aligned}\tag{2}$$

是一致的。

假定分布族(2)也满足上述(i)——(iv)的正则性条件。

下面我们先叙述Wilks 定理

Wilks 定理 设Θ是R_k中的一个有内点的集合，分布族(1)与分布族(2)都满足上述的(i)—(iv)的正则性条件。参数真值$\theta^0\in\Theta_0$且相应于A的内点φ^0。$X_1,\cdots,X_n$为X的iid样本。又设

$$\hat{\theta}_n=\hat{\theta}_n(X_1,\cdots,X_n)$$

和 $$\hat{\varphi}_n=\hat{\varphi}_n(X_1,\cdots,X_n)$$

分别为在分布族(1)和(2)中θ和φ的似然方程的相合解（或称一

致性解)，而

$$LR^*(X_1,\cdots,X_n)\triangleq\frac{\prod_{i=1}^{n}f(X_i;\hat{\theta}_n)}{\prod_{i=1}^{n}\tilde{f}(X_i;\hat{\varphi}_n)},$$

则当$n\to\infty$时，变量

$$Y_n=2\log LR^*(X_1,\cdots,X_n).$$

有极限分布$\chi^2(k-r)$（见〔7〕P327定理3.6.1)。

Wilks 定理是在正则性条件下建立的。当这些正则性条件不成立时，定理的结论也可以不成立。违反这种正则性条件的最重要情况是集合：$\{x: f(x;\theta)>0\}$

与θ有关。我们举一反例来说明这一点。

例1 设$X_1,X_2,\cdots,X_n$为抽自分布族$\{U(0,\theta),\theta>0\}$的 iid 样本。今欲检验假设

$$H_0: \theta=\theta_0(\theta_0>0\text{已知}) \longleftrightarrow H_1: 0<\theta\neq\theta_0$$

(此处表明Θ_0是$r=0$维的)。

易见

$$\sup_{\theta>0}\prod_{i=1}^{n}f(x_i;\theta)=\left(\frac{1}{\max(X_1,\cdots,X_n)}\right)^n.$$

由此得到似然比

$$LR(X_1,\cdots,X_n)=\frac{\sup_{\theta>0}\prod_{i=1}^{n}f(x_i;\theta)}{\sup_{\theta=\theta_0}\prod_{i=1}^{n}f(x_i;\theta)}$$

$$=\frac{[\max(X_1,\cdots,X_n)]^{-n}}{(\theta_0)^{-n}}$$

$$=\left\{\frac{\max(X_1,\cdots,X_n)}{\theta_0}\right\}^{-n}.$$

但当H_0成立时

$$P\left\{\frac{\max(X_1,\cdots,X_n)}{\theta_0}<t\right\}=[P(X<\theta_0 t)]^n$$

$$=\left(\int_0^{\theta_0 t}\frac{1}{\theta}du\right)^n=\left(\frac{\theta_0 t}{\theta}\right)^n,$$

即在 H_0 成立时$\frac{\max(X_1,\cdots,X_n)}{\theta_0}$有密度函数$nt^{n-1}$（$0<t<1$，在此之外为0）。由此得到

$$Y_n=2\ln LR(X_1,\cdots,X_n)=-2n\ln\frac{\max\limits_{1\leqslant i\leqslant n}X_i}{\theta_0}.$$

于是 $$P\left\{-2n\ln\frac{\max\limits_{1\leqslant i\leqslant n}X_i}{\theta_0}<t\right\}$$

$$=P\left\{\ln\frac{\max\limits_{1\leqslant 0\leqslant n}X_i}{\theta_0}>-\frac{t}{2n}\right\}$$

$$=1-P\left\{\frac{\max\limits_{1\leqslant i\leqslant n}X_i}{\theta_0}\leqslant e^{-\frac{t}{2n}}\right\}$$

$$=1-\int_0^{e^{-\frac{t}{2n}}}nu^{n-1}du=1-e^{-\frac{t}{2}}.$$

∴ Y_n之密度函数为

$$\frac{1}{2}e^{-\frac{t}{2}}=\frac{1}{2^{\frac{2}{2}}\Gamma\left(\frac{2}{2}\right)}t^{\frac{2}{2}-1}e^{-\frac{t}{2}}=x^2(2).$$

但此处$K=1$，而$r=0$，故如按 Wilks 定理，极限分布应为$\chi^2(1)$。

10—6 对固定的样本大小而言，似然比检验可以不是无偏的例子

有关定义和定理

定义1 记$X=(X_1,\cdots,X_n)$。设$\phi(X)$为假设检验问题

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta - \Theta_0 \tag{1}$$

的一个检验，其势函数为

$$\beta_\phi(\theta) = E\,\phi(X)。$$

若 $$\beta_\phi(\theta_1) \leqslant \beta_\phi(\theta_2)$$

对任何$\theta_1 \in \Theta_0$及$\theta_2 \in \Theta - \Theta_0$，则称$\phi(X)$为问题（1）的一个无偏检验。

定义2 设对每个样本大小n，确定了问题(1) 的一个检验$\phi_n(X_1, \cdots, X_n)$，以$\beta_n(\theta)$记其势函数，如果存在$\alpha < 1$，使

$$\lim_{n\to\infty} \beta_n(\theta) \leqslant \alpha，\ 对任何\theta \in \Theta_0。 \tag{2}$$

$$\lim_{n\to\infty} \beta_n(\theta) = 1，\ 对任何\theta \in \Theta - \Theta_0。 \tag{3}$$

则称$\phi_n(X)$为问题(1) 的一个相合检验。

我们有如下的结果：

定理 设9—18之Wald定理及10-5之 Wilks 定理的条件成立，而Θ_0为Θ的一个闭子集，则检验问题(1)的似然比检验是相合的。(详见[7]P334定理3.6.2)

从这个似然比检验的相合性定理可知，当样本大小很大时，似然比检验是近似无偏的，这种“渐近无偏性”是任何相合检验都具有的性质。然而，对固定的样本大小而言，似然比检验可以不是无偏的。我们有如下的

例1 设$X_1, X_2, \cdots, X_n$是从母体$N(\mu, \sigma^2)$ 中抽取的 iid 样本，欲检验假设

$$H_0: \sigma^2 = \sigma_0^2 \longleftrightarrow H_1: \sigma^2 \neq \sigma_0^2。 \tag{4}$$

不难算出似然比为

$$LR(X_1, X_2, \cdots, X_n) = \left(\frac{S_n^2}{\sigma_0^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{n}{2}\left(\frac{S_n^2}{\sigma_0^2} - 1\right)\right\},$$

其中 $$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2。$$

若记 $t=\frac{nS_n^2}{\sigma^2}$，

则当 H_0成立时

$$t\sim x^2(n-1).$$

显然，似然比与$\frac{t}{n}e^{-\frac{t}{n}}$（$H_0$成立时）只相差一常数倍。

由 $(xe^{-x})'=e^{-x}(1-x)$

可知函数$xe^{-x}(x>0)$在$x<1$时增加，而在$x>1$时下降。因此，检验问题(4) 的水平为α的似然比检验有否定域

$$\{t:\ t\leqslant d\text{或}t\geqslant b\}. \tag{5}$$

此处d，b由方程组

$$\left.\begin{aligned}&P\{\chi^2(n-1)\leqslant d\}+1-P\{\chi^2(n-1)\leqslant b\}=\alpha,\\&de^{-\frac{d}{n}}=be^{-\frac{b}{n}}\end{aligned}\right\} \tag{6}$$

所决定。

但文献〔7〕中P274例3.3.6已证明了检验问题(4) 的水平为α的一致最优势无偏检验（记为UMPU检验）具有否定域

$$\{t:\ t\leqslant d_1\text{或}t\geqslant b_1\}$$

其中b_1和d_1由方程组

$$\left.\begin{aligned}&P\{\chi^2(n-1)\leqslant d_1\}+1-P\{\chi^2(n-1)\leqslant b_1\}=\alpha\\&d_1^{\frac{n-1}{2}}e^{-\frac{d_1}{2}}=b_1^{\frac{n-1}{2}}e^{-\frac{b_1}{2}}\end{aligned}\right\} \tag{7}$$

所决定。

容易看出，除非$\alpha=1$，否则(6)、(7) 这两个方程组的解必然不同。事实上，若$d=d_1$且$b=b_1$，则(6) 式两边自乘$\frac{n-1}{2}$后，再除以(7)式得

$$\frac{d^{\frac{n-1}{2}}e^{-\frac{d(n-1)}{2n}}}{d^{\frac{n-1}{2}}e^{-\frac{d}{2}}}=\frac{b^{\frac{n-1}{2}}e^{-\frac{b(n-1)}{2n}}}{b^{\frac{n-1}{2}}e^{-\frac{b}{2}}},$$

因而 $e^{\frac{d}{2n}}=e^{\frac{b}{2n}}$，$\frac{d}{2n}\ln e=\frac{b}{2n}\ln e$，

即$b=d$，于是由(6)的第一式得

$$P\{x^2(n-1)\leqslant d\}+1-P\{x^2(n-1)\leqslant d\}=\alpha,$$

得$\alpha=1$。因而，当$\alpha<1$时，(5)式与UMPU检验的否定域不一样。这就证明了：以(5)式为否定域的似然比检验不能是无偏检验。因为，若它是无偏的检验，则如多参数指数族最基本的定理([7]P268定理3.3.5)所证，具有象(6)那样的否定域的无偏检验，必然是UMPU检验。而由UMPU检验的唯一性，将有

$$d=d_1,\ b=b_1.$$

而我们已证明$\alpha<1$时，方程(6)与(7)必然不同，即$d\neq d_1$，$b\neq b_1$。

10—7 一个很坏的似然比检验的例子

有关概念和引理

设$X=(X_1,X_2,\cdots,X_n)$是从母体$\boldsymbol{X}$中抽取的子样，$\boldsymbol{X}$的可能分布族为$\{f(x;\theta),\ \theta\in\Theta\}$，其中$\theta$（可以是向量）是未知参数（当$\boldsymbol{X}$是连续变量时，$f$表示分布密度，当$\boldsymbol{X}$是离散型变量时$f$表示概率分布）。要求检验假设

$$H_0:\ \theta\in\Theta_0\longleftrightarrow H_1:\ \theta\in\Theta-\Theta_0$$

统计量

$$\lambda(x)=\frac{\sup\limits_{\theta\in\Theta}\prod\limits_{i=1}^{n}f(x_i;\theta)}{\sup\limits_{\theta\in\Theta_0}\prod\limits_{i=1}^{n}f(x_i;\theta)}$$

称为似然比。

当 $\lambda(x)>\lambda_0$ (1)

时拒绝H_0。其中λ_0选取使得

$$P_\theta\{\lambda(x)>\lambda_0\}\leqslant\alpha \quad \text{对一切}\theta\in\Theta_0。\tag{2}$$

由(1)、(2) 两式给出的检验法称为水平为α的似然比检验。

对于简单假设

$$H_0:\ \theta=\theta_0 \longleftrightarrow H_1:\ \theta=\theta_1\ (\theta_1\neq\theta_0).$$

我们有下面著名的Neyman-Pearson 基本引理：

Neyman-Pearson 基本引理　设子样$X=(X_1,\cdots,X_n)$是连续型随机向量，分布密度是$f(x;\theta)$，θ是未知参数．对于简单假设检验问题：

$$H_0:\ \theta=\theta_0 \longleftrightarrow H_1:\ \theta=\theta_1\ (\theta_1\neq\theta_0).$$

若存在检验函数$\phi(X)$使

$$E_{\theta_0}\phi(X)=\alpha,\tag{3}$$

$$\phi(x)=\begin{cases}1, & \dfrac{f(x;\theta_1)}{f(x;\theta_0)}>K,\\[2ex] 0, & \dfrac{f(x;\theta_1)}{f(x;\theta_0)}<K\ (K\text{ 为大于等于零的常数})。\end{cases}\tag{4}$$

成立的必要充分条件为$\phi(X)$是水平为α的最优势检验(记为MP检验)。

这引理说明，对于简单对简单的假设检验问题，似然比检验法是最优的；反之，最优势检验法一定是似然比检验法．但是，对于某些简单对复杂的假设检验问题，似然比检验有时可以给出很坏的结果．下面的例子是Stein和 Rubin给出的。

例1　设X是取0、±1、±2 这五个点的离散型随机变量，其分布

$$
P_{p=0}(X=x)=\begin{cases}\dfrac{a}{2}, & \text{当}x=\pm 2\text{时},\\[2mm] \dfrac{1-2a}{2}, & \text{当}x=\pm 1\text{时},\\[2mm] a, & \text{当}x=0\text{时},\end{cases}
$$

$$
P_p(X=x)=\begin{cases}pc, & \text{当}x=-2,\\ (1-p)c, & \text{当}x=2,\\ \alpha\left(\dfrac{1-c}{1-a}\right), & \text{当}x=0\quad(0<p<1),\\[2mm] \dfrac{1-c}{1-a}\left(\dfrac{1}{2}-a\right), & \text{当}x=\pm 1.\end{cases}
$$

其中a、c为常数，满足

$$
0<a<\frac{1}{2}
$$

和
$$
\frac{a}{2-a}<c<a.
$$

现要根据对X的一次观测，求假设检验问题

$$
H_0:\ p=0\longleftrightarrow H_1:\ p\neq 0 \tag{*}
$$

的水平为a的似然比检验。

下面先来计算似然比

$$
\lambda(2)=\frac{\sup\limits_{0<p<1}P_p(\chi=2)}{P_{p=0}(\chi=2)}=\frac{\sup\limits_{0<p<1}\left\{\dfrac{\alpha}{2},(1-P)c\right\}}{\dfrac{\alpha}{2}}
$$

$$
=\frac{c}{\dfrac{a}{2}}=\frac{2c}{a}
$$

$$
\left(\because\quad c>\frac{a}{2-a}>\frac{\alpha}{2}\right),
$$

$$\lambda(-2)=\frac{\sup\limits_{0\leq p<1}P_p(x=-2)}{P_{p=0}(x=-2)}=\frac{2c}{a},$$

$$\lambda(1)=\lambda(-1)=\frac{\sup\limits_{0\leq p<1}P_p(x=\pm 1)}{P_{p=0}(x=\pm 1)}=\frac{1-c}{1-a},$$

$$\lambda(0)=\frac{\sup\limits_{0\leq p<1}P_p(x=0)}{P_{p=0}(x=0)}=\frac{1-c}{1-a}.$$

即

$$\lambda(x)=\begin{cases}\dfrac{2c}{a}, & \text{当}x=\pm 2\text{时},\\[2ex] \dfrac{1-c}{1-a}, & \text{当}x=0,\pm 1\text{时}.\end{cases}$$

∵

$$P_{p=0}\left(\lambda(X)=\frac{1-c}{1-a}\right)=P_{p=0}(x=0)+P_{p=0}(x=1)+P_{p=0}(x=-1)$$

$$=a+\left(\frac{1}{2}-a\right)+\left(\frac{1}{2}-a\right)=1-a,$$

$$P_{p=0}\left(\lambda(X)=\frac{2c}{a}\right)=P_{p=0}(x=+2)+P_{p=0}(x=-2)$$

$$=\frac{a}{2}+\frac{a}{2}=a,$$

$$P_{p\neq 0}\left(\lambda(X)=\frac{1-c}{1-a}\right)=P_{p\neq 0}(x=0)+P_{p\neq 0}(x=1)+P_{p\neq 0}(x=-1)$$

$$=\frac{1-c}{1-a}a+2\cdot\frac{1-c}{1-a}\left(\frac{1}{2}-a\right)$$

$$=\frac{1-c}{1-a}\left[a+2\left(\frac{1}{2}-a\right)\right]=1-c,$$

$$P_{p\neq 0}\left(\lambda(X)=\frac{2c}{a}\right)=P_{p\neq 0}(x=2)+P_{p\neq 0}(x=-2)$$

$$= (1-P)c + pc = c,$$

所以当 H_0: $p=0$, H_1: $p\neq 0$成立时

(X)的分布为

λ 取 值	$\frac{1-c}{1-a}$	$\frac{2c}{a}$
H_0: $p=0$	$1-a$	a
H_1: $p\neq 0$	$1-c$	c

对于$0<a<1$，我们要决定λ_0使得

$$P_{p=0}(\lambda(X)\geqslant\lambda_0)\leqslant a.$$

$\because$ $\quad \frac{a}{2-a}<c,$

$\therefore$ $\quad a<2c-ac$

于是 $\quad a-ca<2c-ac-ac,$

即 $\quad a(1-c)<2c(1-a),$

$\therefore$ $\quad \frac{a}{2c}<\frac{1-a}{1-c},\ \frac{1-c}{1-a}<\frac{2c}{a}.$

故取λ_0为

$$\frac{1-c}{1-a}<\lambda_0<\frac{2c}{a},$$

则取似然比检验$\phi_1(\chi)$：当$\lambda\geqslant\lambda_0$时，否定$H_0$。就满足要求，即

$$\phi_1(x)=\begin{cases}1, & 当\lambda(x)\geqslant\lambda_0,\\ 0, & 当\lambda(x)<\lambda_0\end{cases}$$

就满足要求。这是因为：

$$P_{p=0}(\lambda(X)\geqslant\lambda_0)=P_{p=0}\left(\lambda(x)=\frac{2c}{a}\right)=P_{p=0}(x=\pm 2)$$

$$=a.$$

而 $\quad P_{p\neq 0}(\lambda(X)\geqslant\lambda_0)=P_{p\neq 0}\left(\lambda(X)=\frac{2c}{a}\right)$

$$=P_{p\neq 0}(x=\pm 2)=c<\alpha,$$

即似然比检验$\phi_1(X)$在$p\neq 0$处的势函数比不作任何试验，而径取$\phi(X)\equiv\alpha$还不如。

更有甚者，我们可以构造一个检验，它与问题（*）毫无关系，但却比检验$\phi_1(X)$好，这个检验是：设α为有理数，取适当大的N，使$N\alpha$、$N(1-\alpha)$为正整数，在装有$N\alpha$个红球，$N(1-\alpha)$个白球的袋中，随机地取出一个球，若为红球，就否定H_0。（记这个检验为$\phi_2(X)$）。

因为每取一次，取出红球的概率为$\frac{N\alpha}{N}=\alpha$，若摸出红球，就否定H_0，则有

$$P_{p=0}(\text{否定}H_0)=\alpha,$$

$$P_{p\neq 0}(\text{否定}H_0)=\alpha>c=P_{p\neq 0}(\lambda(X)\geqslant\lambda_0),$$

即检验$\phi_2(X)$的水平为α，势函数也为α，它当然比检验$\phi_1(X)$要好。

再看另一个检验$\phi_3(X)$：当$x=0$时，否定H_0。对这个检验，有

$P_{p=0}(x=0)=\alpha$，$P_{p\neq 0}(x=0)=\frac{1-c}{1-\alpha}\alpha>\alpha$，所以也比检验$\phi_1(X)$要好。从下表

X	-2	2	0	± 1
$\phi_1(X)$	否定H_0	否定H_0	接受H_0	接受H_0
$\phi_3(X)$	接受H_0	接受H_0	否定H_0	接受H_0

可以看出，$\phi_1(X)$与$\phi_3(X)$除了在点± 1处相一致外，在其余的点上都是反其道而行之，但是$\phi_3(X)$却比似然比检验$\phi_1(X)$要好。之所以会出现这种情况，这是因为似然比检验纯粹是一

种直观的技术，它并不是基于一种明确的优良性考虑。因此，虽然在许多问题上常给出具有优良性质的检验，但在个别情况下，仍然可能给出很坏的检验。

第十一章　与线性模型有关的反例

线性模型是一种在理论上和应用上都很重要的统计模型。关于它我们在这里只提供以下一个很重要的反例。

11—1　在非正态的条件下，参数的最小方差线性无偏估计不再是最小方差无偏估计的例子

有关定理

关于线性模型参数的最小二乘估计，我们有以下著名的Gauss—Markov 定理：

定理1 (Gauss—Markov)：

设线性模型为：$(Y,\ X\beta,\ \sigma^2 In)$

$$\hat{\beta}_L=(X'X)^{-1}X'Y$$

是参数β的最小二乘估计，则$\hat{\beta}_L$是β的最小方差线性无偏估计。(见〔6〕第二册之一分册P93)

定理1断言β的最小二乘估计$\hat{\beta}_L$在一切线性无偏估计中是方差最小的。但是，它并没有排斥这种可能性，即存在方差比$\hat{\beta}_L$更小的非线性无偏估计。但是，如果我们进一步假定

$$Y\sim N(X\beta,\ \sigma^2 In),$$

则这种可能性不存在，即我们有以下的结论：

定理2 在观察向量$Y \sim N(X\beta, \sigma^2 In)$的条件下，线性模型$(Y, X\beta, \sigma^2 In)$中参数$\beta$的最小方差线性无偏估计

$$\hat{\beta}_L = (X^\tau X)^{-1} X^\tau Y$$

也是β的最小方差无偏估计。

证明 $Y \sim N(X\beta, \sigma^2 In)$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}[(Y-X\beta)^\tau (Y-X\beta)]}$$

对任一Y的函数$f(Y)$（不必是Y的线性函数）设其属于集合：

$$U_0 = \{\nu: E_\beta \nu = 0, E_\beta \nu^2 < \infty, \text{对一切}\beta \in \Theta\},$$

则 $$Ef(Y) = \int f(Y)\, e^{-\frac{1}{2\sigma^2}[(Y-X\beta)^\tau (Y-X\beta)]} dY = 0,$$

即 $$\int f(Y)\, e^{-\frac{1}{2\sigma^2}[Y^\tau Y + \beta^\tau X^\tau X\beta - 2\beta^\tau X^\tau Y]} dY = 0.$$

由此有 $$\int f(Y)\, e^{-\frac{Y^\tau Y}{2\sigma^2} + \frac{2\beta^\tau X^\tau Y}{2\sigma^2}} dY = 0,$$

即 $$\int f(Y) e^{-\frac{Y^\tau Y}{2\sigma^2} + \frac{(X^\tau Y)^\tau \beta}{\sigma^2}} dY = 0.$$

上式关于β微分得

$$\int f(Y) e^{-\frac{Y^\tau Y}{2\sigma^2} + \frac{(X^\tau Y)^\tau \beta}{\sigma^2}} \frac{1}{\sigma^2} X^\tau Y dY = 0,$$

即有 $$\int (X^\tau Y) f(Y) e^{-\frac{(Y-X\beta)^\tau (Y-X\beta)}{2\sigma^2}} dY = 0.$$

也即 $$E(X^\tau Y f(Y)) = 0.$$

于是由〔6〕第二册之二分册P150定理1知 $X^\tau Y$ 是其数学期望$EX^\tau Y = X^\tau X\beta$的最小方差无偏估计，再由〔6〕第二册之二分册P152之系1知 $(X^\tau Y)$的线性函数$(X^\tau X)^{-1} X^\tau Y$是其数学期望

$$E(X^\tau X)^{-1} X^\tau Y = E\hat{\beta}_L = \beta$$

的最小方差无偏估计，证毕。

但是定理2的逆不真，我们有如下的反例：

例1 设

$$\begin{cases} Y_1 = -\beta_1 + \beta_2 + x_{13}\beta_3 + e_1, \\ Y_2 = x_{21}\beta_1 + x_{22}\beta_2 + x_{23}\beta_3 + e_2, \\ Y_3 = x_{31}\beta_1 + x_{32}\beta_2 + x_{33}\beta_3 + e_3, \end{cases}$$

其中 $e_i \sim U[-2\sqrt{3}\sigma,\ 2\sqrt{3}\sigma]$，$i=1,\ 2,\ 3$，

且e_i之间相互独立（$i=1,\ 2,\ 3$）。则

$$\begin{cases} Y_1 \sim U[-\beta_1 + \beta_2 + x_{13}\beta_3 - 2\sqrt{3}\sigma,\ -\beta_1 + \beta_2 \\ \quad + x_{13}\beta_3 + 2\sqrt{3}\sigma] \triangleq U[a_1,\ b_1], \\ Y_2 \sim U[x_{21}\beta_1 + x_{22}\beta_2 + x_{23}\beta_3 - 2\sqrt{3}\sigma,\ x_{21}\beta_1 \\ \quad + x_{22}\beta_2 + x_{23}\beta_3 + 2\sqrt{3}\sigma] \triangleq U[a_2,\ b_2], \\ Y_3 \sim U[x_{31}\beta_1 + x_{32}\beta_2 + x_{33}\beta_3 - 2\sqrt{3}\sigma,\ x_{31}\beta_1 \\ \quad + x_{32}\beta_2 + x_{33}\beta_3 + 2\sqrt{3}\sigma] \triangleq U[a_3,\ b_3]. \end{cases}$$

由定理1知

$$\hat{\beta}_L = (X^\tau X)^{-1} X^\tau Y$$

是$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$的最小方差线性无偏估计。

其中 $X = \begin{pmatrix} -1 & 1 & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$，$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$。

现在我们来证明$\hat{\beta}_L$不是β的最小方差无偏估计。

记 $U_0 \triangleq \{\nu:\ E_\beta \nu = 0,\ E_\beta \nu^2 < \infty$，对一切$\beta \in Ⓗ\}$，

$U \triangleq \{T:\ E_\beta T = \beta,\ E_\beta T^2 < \infty$，对一切$\beta \in Ⓗ\}$。

由〔6〕第二册之二分册p150之定理1知$\hat{\beta}_L$是β 的最小方差无偏估计之充要条件为

$$E_\beta(\nu\hat{\beta}_L) = 0，对一切\beta \in Ⓗ和\nu \in U_0。$$

为此，我们只需证明，存在$\beta_0 \in Ⓗ$和$\nu_0 \in U_0$使

$$E_{\beta_0}(\nu_0\hat{\beta}_L)\neq 0$$

就够了．取

$$\beta_0=\begin{pmatrix}1\\1\\0\end{pmatrix},\quad \nu_0=Y_1\in U_0\quad(\because E_{\beta_0}Y_1=0).$$

则
$$E_{\beta_0}(\nu_0\hat{\beta}_L)=E_{\beta_0}(Y_1(X_\tau X)^{-1}X^\tau Y)$$
$$\neq 0\Longleftrightarrow (X^\tau X)^{-1}E_{\beta_0}(Y_1X^\tau Y)$$
$$\neq 0\Longleftrightarrow E_{\beta_0}(Y_1X^\tau Y)$$
$$\neq 0\Longleftrightarrow E_{\beta_0}\left[Y_1\begin{pmatrix}\sum_{i=1}^{3}X_{i1}Y_i\\ \sum_{i=1}^{3}X_{i2}Y_i\\ \sum_{i=1}^{3}X_{i3}Y_i\end{pmatrix}\right]$$
$$=\begin{pmatrix}E_{\beta_0}\left(Y_1\sum_{i=1}^{3}X_{i1}Y_i\right)\\ E_{\beta_0}\left(Y_1\sum_{i=1}^{3}X_{i2}Y_i\right)\\ E_{\beta_0}\left(Y_1\sum_{i=1}^{3}X_{i3}Y_i\right)\end{pmatrix}\neq 0\quad(其中x_{11}=-1,\ x_{12}=1),$$

而
$$E_{\beta_0}\left(Y_1\sum_{i=1}^{3}X_{i2}Y_i\right)$$
$$=\iiint Y_1(X_{12}Y_1+X_{22}Y_2+X_{32}Y_2)\prod_{i=1}^{3}\frac{1}{b_i-a_i}dY_1dY_2dY_3$$
$$=C\iiint(X_{12}Y_1^2+X_{22}Y_1Y_2+X_{32}Y_1Y_3)dY_1dY_2dY_3$$

$$= CX_{12}E_{\beta_0}Y_1^2 + CX_{22}E_{\beta_0}Y_1E_{\beta_0}Y_2 + CX_3E_{32}Y_1E_{\beta_0}Y_3$$

$$= CX_{12}E_{\beta_0}Y_1^2 = CX_{12}[D_{\beta_0}Y_1 + (E_{\beta_0}Y_1)^2]$$

$$= CX_{12}\left[\frac{(b_1 - a_1)^2}{12}\right] = CX_{12}\left[\frac{(4\sqrt{3}\sigma)^2}{12}\right]$$

$$= CX_{12} \cdot 4\sigma^2 \neq 0,$$

其中 $$C = \sum_{i=1}^{3} \frac{1}{b_i - a_i}.$$

证毕.

注 定理2的证明也可参看〔7〕p451之定理5.2.4.

第十二章　与抽样理论有关的反例

12—1　当检查人员可能犯错误时，全面检查不一定比抽样检查更好的例子

本章我们介绍波兰应用数学工作者得到的一个结果：即当检查人员可能犯错误时，全面检查不一定比抽样检查更好的反例。

我们知道，全面的验收检查所制订的验收计划常常是这样的：检查全部母体，坏的个体拒收，好的则接受。若检查人员不犯错误，且如不顾及检查费用时，则全面检查这种方法是最好的，因为它保证了一切次品的被拒收，故经过挑选而接受的母体中其次品数为零。

但如果检查人员可能犯错误，则虽然全部检查，仍不能保证把次品全部挑出，且此外还可能把某些上品错误地抛掉。这可从Kennedy和Mierzejewska所做的试验中得到证实。Kennedy的试验是：在一批全为上品的母体(其个数未知)中，添进100个次品(假定个体的类别只有上品或次品两种)，然后把这批产品拿去检查，结果检查者只找出了68个次品，剩下来的总体请他再检查一次，而不告诉他这就是他先前已检查过的且经他认为全是上品的总体。在第二次分类中，他又找到了18个次品，再让他作第三次检查，又发现了8个。第四次特别组织3人去检查，

又找到了4个次品。但还有2个次品未找出来。Mierzejewska所做的一次试验是：有1873个小齿轮，检查者在分类时挑出了169个次品后，她又去检查了几次。第一次她找到了70个次品，第二次24个，第三次4个，第四次1个次品。然后，她再一次检查了那些认为是次品的小齿轮，结果发现在第一次所挑出的70个中有2个是上品，第二次的24个中有1个也是上品。这些例子表明，在某些技术条件下，检查人员犯错误（误认上品为次品的第一类错误及误认次品为上品的第二类错误）的概率不会太小，如不注意这一点，就可能得出本质上错误的结论。同样的情形不仅在品质检查中，而且在排版校对、X光照像底片的阅读，以及其他场合遇到。

下面，我们在某些条件下，在估计到检查人员的错误时，抽样检查给出比全面检查更好的例子。这个例子不仅很有趣而且很有实用价值。

例1 假定在接受某一个总体时，检查人员以概率$1-p$认上品为次品，以概率$1-q$认次品为上品（因而p及q为得到正确结论的概率：认上品为上品及认次品为次品）。这里要注意概率p及q依赖于许多难以细致研究的检查条件（如检查人员的工作速度；受检查的样本或总体的个数；检查人员的疲劳程度；总体的次品率等）。因此，它们甚至在检查一批总体也会显著地变化。

如已知总体的次品率为w，则从总体中随机地挑出的个体被认为是上品的概率为

$$P\{C|w,\ p,\ q\}=(1-w)p+w(1-q),$$

因而，从总体中随机地挑出的个体被认为是次品的概率

$$P\{E|w,\ p,\ q\}=1-P\{C|w,\ p,\ q\},$$

$$\begin{aligned}P\{E|w,\ p,\ q\}&=1-[(1-w)p+w(1-q)]\\&=1-(1-w)p-w(1-q)\end{aligned}$$

$$= wq + (1-w)(1-p) \triangleq w_1.$$

后一概率 w_1 为总体的伪次品率，因为运用全面检查时被认为是次品而拒收的个体的频率就是 w_1。

下面，我们来给出检查人员不常正确时，抽样检查比全面检查更好而应满足的条件。

设有个体数为 N，次品率为 w 的总体，进行全检查后恰有 k 个个体被认为是次品的概率 $P_k(N, w, p, q)$ 为

$$P_k(N,w,p,q) = \sum_{i=0}^{k} \binom{N(1-w)}{i} \binom{Nw}{k-i} p^{N(1-w)-i} (1-p)^i q^{k-i} \cdot (1-q)^{Nw-(k-i)} \tag{1}$$

其中因子 $\binom{N(1-w)}{i} p^{N(1-w)-i}(1-p)^i$

为总体中 $N(1-w)$ 个上品的个体里，恰有 i 个被认为是次品的概率，而另一因子

$$\binom{Nw}{k-i} q^{k-i}(1-q)^{Nw-(k-i)}$$

为总体中 Nw 个次品的个体里，有 $k-i$ 个被认为是次品的概率。

公式中的求和取自一切可能的 i，即自0到 k。

今考虑验收方案 $m \mathbin{/\!/} n$（或记为 (n, m)）的抽样检查。记号 $m \mathbin{/\!/} n$ 表示从总体中抽取 n 个样本，若其中的次品数小于等于 m，则接受这批产品。如检查人员不犯错误，则方案 $m \mathbin{/\!/} n$ 的特性函数由下式计算

$$f(w) = P\{z \leqslant m\} = \sum_{i=0}^{m} \frac{\binom{Nw}{i}\binom{N(1-w)}{n-i}}{\binom{N}{n}}.$$

如检查人员可能犯错误，则方案m/n的特性函数为

$$\varphi(w)=P(k\leqslant m)$$

$$=\sum_{j=0}^{m}\sum_{i=0}^{n}\frac{\binom{Nw}{i}\binom{N(1-w)}{n-i}}{\binom{N}{n}}P_j\left(n,\ \frac{i}{n},\ p,q\right),\qquad (2)$$

其中$\dfrac{\binom{Nw}{i}\binom{N(1-w)}{n-i}}{\binom{N}{n}}$为下列事件的概率：自个数为$N$，次品率为$w$的总体中，抽取个数为$n$的样本，此样本中恰含$i$个次品。

而$P_j\left(n,\ \frac{i}{n},\ p,\ q\right)$是检查人员检查此样本后，其中恰有$j$件被认为是次品的概率，它由(1)式表出。

如$k\leqslant m$，则应用抽样检查时，总体全被接受。这等于把未列入样本中的一切个体划为上品。

如$k>m$，则总体被拒收，这等于承认全体未经检查的个体都是次品。

我们现在来比较全面检查法和抽样检查法中，被错误地分类的个体的数学期望。

在全面检查法中，被错误地分类的个体件数的数学期望R_1为

$$R_1=N[1-(1-w)p-wq].\qquad (3)$$

如采用抽样检查方案，则在样本中被错误地分类的个体件数的数学期望为

$$\tau_1=n[1-(1-w)p-wq].\qquad (4)$$

如果根据抽出来的样本，总体被拒收，则其他一切个体都

算是次品．因为总体被拒收的概率为$1-\varphi(w)$，故未经检查的个体中，原系上品而被认为是次品的件数的数学期望为

$$\tau_2=(N-n)(1-w)[1-\varphi(w)]. \tag{5}$$

同样，根据验收方案 $m/\!\!/n$，总体被接受时，未经检查的个体中，原系次品而被认为是上品的件数的数学期望为

$$\tau_3=(N-n)w\varphi(w). \tag{6}$$

把τ_1，τ_2，τ_3加起来，就得到抽样方法中被错误地分类的个体件数的数学期望R_2

$$\begin{aligned}R_2&=n[1-(1-w)p-wq]+(N-n)(1-w)[1-\varphi(w)]\\&\quad+(N-n)w\varphi(w)\\&=n[1-(1-w)p-wq]+(N-n)[1-w\\&\quad-(1-2w)\varphi(w)].\end{aligned} \tag{7}$$

比较公式(3)与(7)，可知，当下面的不等式成立时，抽样方法优于全面检查($R_2\leqslant R_1$)

$$(1-2w)\varphi(w)\geqslant p(1-w)-w(1-q). \tag{8}$$

$\because\varphi(w)\geqslant0$，故如上面不等式的右方取负值时(8)式成立。

为此，只要当$1-2w>0$时

$$\frac{p}{1+p-q}<w<\frac{1}{2},$$

如
$$\frac{p}{1+p-q}<\frac{1}{2}.$$

如$1-2w<0$，由于$\varphi(w)\leqslant1$得

$$\frac{1}{2}<w<\frac{1-p}{1+q-p},$$

如
$$\frac{1-p}{1+q-p}>\frac{1}{2}. \tag{9}$$

(9) 中不等式表明，在某些情况下，抽样检查优于对总体的全

面检查。

这一结果有很大的实际意义，第一，运用抽样检查可以节省许多费用，因此抽样检查几乎总是比全面检查便宜。第二，不检查总体，而只检查样本，在许多情况下，可以把工作做得更好，因而缩小犯错误的概率$1-p$及$1-q$。第三，特别地，当信息具有很强的时间性时，旷日持久的全面检查将只能获得陈旧的信息，而变得毫无价值，因而必须进行抽样检查。

主要参考文献

〔1〕Laha. R. G. and Rohatgi. V. K; Probability theory, 1979年.

〔2〕Rohatgi. V. K: An introduction to probability theory and Mathematical statistics, 1976年。

〔3〕M. Loéve: Probability theory (1) 4th edition, 1977. (第三版有中译本)

〔4〕И. И Гихман等; Теория Вероятностей и математическая статистика.

〔5〕王梓坤; «概率论基础及其应用», 科学出版社, 1976.

〔6〕复旦大学数学系;«概率论»第一、二册, 人民教育出版社.1982

〔7〕陈希孺; «数理统计引论», 科学出版社,1981年.

〔8〕D. A. S. Fraser; Nonparametric methods in statistics, 1956.

〔9〕P. J. Bickel, K. A. Doksum; Mathematical statistics Basic ideas and seleted topics Halden-Day inc, 1977.

〔10〕N. C. Giri, Introduction to Probability and statistics Part II: statistics. Mareel Dekker inc 1975

〔11〕K. L Chung; A Course in Probability Theory, Harconrt, Brace and World, New York, 1968.

〔12〕Steven F. Arnald; The theory of linear Models and Multivariate Analysis, John wiley and sons, 1980.

〔13〕W. Feller; «概率论及其应用» (上册) 1950年版, 科学出版社.

〔14〕杨宗磐«概率论入门», 1981, 科学出版社.

〔15〕Halmos. P. R; Measure Theory, 1950.

〔16〕H. Cramer; Mathematical Methods of statistics, 1946. (有汉译本, 魏宗舒译.统计学数学方法)

〔17〕S. S. Wilks, Mathematical Statistics, 1962.

〔18〕Б. В. Гнеденко, 概率论教程 第二版 1954. (丁寿田译)

〔19〕Y. S. Chow; H. Teicher; Probability Theory, 1978.

[General Information]
书名=概率统计中的反例
作者=张尚志　　刘锦萼编著
页数=312
SS号=11179618
出版日期=1988年03月第1版

前言
目录
目录
第一章　与随机事件、概率空间、古典概型有关的反例
1－1　基本事件但不是事件的例子
1－2　同一随机现象可以用不同的样本空间来描述的例子
1－3　两事件互斥但不互逆的例子
1－4　概率为０的事件并非是不可能事件的例子
1－5　概率为１的事件并非是必然事件的例子
（＊）1－6　集类Ｅ上的非负集函数μ具有有限可加性、连续性但不具有可列可加性的例子
1－7　上极限事件和下极限事件不相等的例子
1－8　非古典型随机试验的例子
第二章　与独立性、条件概率有关的反例
2－1　两事件不独立的例子
2－2　两事件不独立，但不一定互斥的例子
2－3　事件两两独立，但并非相互独立的例子
2－4　Ｐ（ＡＢＣ）＝Ｐ（Ａ）Ｐ（Ｂ）Ｐ（Ｃ），但Ａ，Ｂ，Ｃ三事件并非两两独立的例子
2－5　说明Ｐ（Ａ），Ｐ（Ａ｜Ｂ），Ｐ（ＡＢ）三者不同含义的例子
2－6　非独立试验概型（非Ｂｅｒｎｏｕｌｌｉ概型）的例子
（＊）2－7　？ｓｕｐＥ｜ξｎ｜＜∞，但ξｎ？ξ不成立的例子
（＊）2－8　并非任何一个鞅｛ξｎ｝均存在一个满足Ｅ｜ξ｜＜∞的随机变量ξ和概率空间（Ω，？１，Ｐ）中？１的子σ代数？ｎ使得ξｎ＝Ｅ｛ξ｜？ｎ｝的例子
（＊）2－9　ξ与ζ独立，但Ｅ（ξ｜η，ζ）？Ｅ（ξ｜η）的例子
第三章　与随机变量，分布函数有关的反例
3－1　ξ（ｗ）是定义在概率空间（Ω，？，Ｐ）上的单值实函数，但它不是（Ω，？１，Ｐ）上的随机变量的例子
（＊）3－2　随机变量的勒贝格可测函数不一定是随机变量的例子
3－3　既非离散型，又非连续型的分布函数的例子
3－4　设ξ是一个连续型随机变量，ｇ是某个连续函数，η＝ｇ（ξ）不是连续型随机变量的例子
3－5　不同的随机变量（向量），具有相同的分布

函数的例子

3－6　连续型随机变量之密度函数未必是连续的例子

3－7　二元函数F（x，y）对每个变元非降，左连续且F（x，-∞）＝0，F（-∞，y）＝0，F（+∞，+∞）＝1但仍不是分布函数的例子

3－8　边际分布是正态分布，但联合分布不是多元正态分布的例子

3－9　随机变量ξ、η相互独立，而且同分布，但不一定有ξ＝η（α，s）成立的例子

3－10　ξ、η同分布但不独立时，ξ＝ξ－η不一定是对称随机变量的例子

（＊）3－11　ξ、η服从正态分布，但ξ、η不独立，则ξ＋η不一定服从正态分布的例子

3－12　两个不同的联合分布（函数），它们可以有相同的边际分布的例子

3－13　随机变量ξ1，ξ2，ξ3，两两独立，但不相互独立的例子

3－14　ξ、η不独立，但ξ2和η2独立的例子

3－15　相同的随机向量构造的不同的Borel可测函数（不恒等于常数）之间也可能是独立的例子。

3－16　从随机向量ξ＝（ξ1，ξ2，…，ξm）和η＝（η1，η2，…，ηn）之间的独立性推不出ξ的分量ξ1，ξ2，…，ξm或者η的分量η1，η2，…，ηn之间的独立性的例子

3－17　ξ和η1独立，ξ和η2独立，但ξ和随机向量（η1，η2）不独立的例子

（＊）3－18　非独立随机变量序列的例子

（＊）3－19　分布函数F1和F2的卷积绝对连续，但F1和F

不绝对连续的例子

（＊）3－20　随机变量ξ、η的各阶矩不全存在，即使ξ、η已经存在的各阶矩相等，亦不能推出ξ，η的分布相同的例子

（＊）3－21　ξ与？有相同的分布，而ξ并非服从哥西分布的例子

（＊）3－22　ξ与1－ξ有相同的分布，而ξ并非服从B（α，α）分布的例子

（＊）3－23　ξ与η独立同分布，ξ＝？～？（1，0）＝？，但ξ，η并非服从正态分布的例子

（＊）3－24　ξ与η独立，分别服从B（α1，β1）及B（α2，β

分布，又ξη～B（α，β），则β＝β1＋β2，但

只有α＝α１或α＝α２的例子
（＊）３－２５　？Ｓｋ（ｍｋ＝Ｅξｋ）对Ｓ不绝对收敛，｛ｍｋ｝仍唯一决定ξ之分布函数的例子
第四章　与数字特征有关的反例
４－１　随机变量的数学期望不存在（从而方差也不存在）的例子
４－２　随机变量的数学期望存在，但方差不存在的例子
４－３　任何阶矩都不存在的随机变量的例子
４－４　随机变量ξ的一阶矩存在，但没有更高整数阶矩的例子
４－５　随机变量ξ在０＜ｒ＜１时，Ｅξ存在；但在ｒ≥
时，Ｅ？不存在的例子
４－６　随机变量ξ１和ξ２，它们的一切整数阶矩都相同（即Ｅξ？＝Ｅξ？，ｋ＝１，２，３…），但它们的分布函数不相等的例子
４－７　随机变量ξ１和ξ２不相关，但也不独立的例子
４－８　相关系数ρ（ξ１，ξ２）＞０，ρ（ξ２，ξ３）＞０，但ρ（ζ１，ζ３）＜０的例子
４－９　ξ１和ξ２不独立，但Ｅ（ξ１·ξ２）＝Ｅξ１·Ｅξ２的例子
４－１０　ξ１，ξ２独立，但Ｅ（ξ１＋ξ２）ｋ？Ｅξ？＋Ｅξ？的例子（正整数ｋ？１）
４－１１　Ｅ［Ｅ（η｜ξ）］存在，但Ｅη不存在，因而Ｅη＝Ｅ［Ｅ（η｜ζ）］不成立的例子
４－１２　中位数不唯一的例子
４－１３　数学期望不存在，但中位数存在的随机变量的例子
４－１４　随机变量的众数不唯一的例子
第五章　与特征函数、母函数有关的反例
５－１　随机变量ξ１，ξ２不独立，但？ξ１＋ξ２（ｔ）＝？ξ１（ｔ）·？ξ２（ｔ）成立的例子
５－２　当ｋ为奇数时，随机变量ξ的特征函数？（ｔ）在ｔ＝０处可微分ｋ次，但Ｅξｋ不存在的例子
５－３　分布函数绝对连续，但其对应的特征函数不绝对可积的例子
５－４　特征函数？（ｔ）在有限区间内的值不足以唯一确定此？（ｔ），从而也不足以唯一决定分布函数Ｆ（ｘ）的例子
５－５　特征函数列的极限函数不是特征函数的例子
５－６　分布函数不具有再生性的例子

５－７　　分布函数Ｆ（ｘ）不是无穷可分分布的例子
５－８　　无处为０的特征函数不是无穷可分的例子
５－９　　无穷可分的特征函数可以分解为不是无穷可分的特征函数的乘积的例子
５－１０　　随机变量的矩母函数不存在的例子
５－１１　　随机变量ξ的各阶矩都存在，但矩母函数不存在的例子
（＊）５－１２　　并非所有的特征函数都是解析的例子
（＊）５－１３　　消去法对一般特征函数之分解不一定成立的例子
（＊）５－１４　　无穷可分的特征函数存在不可分解因子的例子
第六章　　与收敛性有关的反例
６－１　　分布函数列｛Ｆｎ（ｘ）｝弱收敛于Ｆ（ｘ），但Ｆ（ｘ）不是分布函数的例子
６－２　　分布函数列Ｆｎ（ｘ）？Ｆ（ｘ），但Ｆ（ｘ）不唯一的例子
６－３　　即使？ξｎ（ｗ）＝ξ（ｗ）对？∈Ω都成立，也不能保证？Ｆｎ（ｘ）＝Ｆ（ｘ）对？ｘ∈Ｒ１成立的例子
６－４　　Ｆｎ（ｘ）？Ｆ（ｘ），但Ｅξ？→Ｅξｋ不成立的例子
６－５　　Ｆｎ（ｘ）？Ｆ（ｘ），但Ｆｎ（ｘ）？Ｆ（ｘ）不成立的例子
６－６　　设分布函数列｛Ｆｎ（ｘ）｝对应的特征函数列为｛？ｎ（ｔ）｝，若｛？ｎ（ｔ）｝收敛于某函数？（ｔ），但？（ｔ）在ｔ＝０不连续，则推不出｛Ｆｎ（ｘ）｝弱收敛于某分布函数的例子
６－７　　由ξｎ？ξ推不出相应的分布密度函数或概率分布的收敛性的例子
６－８　　ξｎ（ｗ）？ξ（ｗ），但ξｎ（ｗ）？ξ（ｗ）不成立的例子
６－９　　ξｎ（ｗ）？ξ（ｗ）但Ｅξ？→Ｅξｋ（？ｋ≥１）不成立的例子
６－１０　　ξｎ？ξ但ξｎ？ξ不成立的例子
６－１１　　波雷尔——康特立引理（１）的逆不成立的例子
６－１２　　ξｎ？ξ但ξｎ？ξ不成立的例子
６－１３　　设０＜ｓ＜ｒ，ξｎ？ξ但ξｎ？ξ不成立的例子
６－１４　　有关ｒ一阶收敛与几乎处处收敛之间关系的反例

（＊）６—１５　　当Ｆｎ？Ｆ时，？Ｆ的例子
（＊）６—１６　　ξｎ依分布收敛到ξ（记为ξｎ？ξ），但ξｎ之矩母函数Ｍｎ（ｓ）不收敛到ξ之矩母函数Ｍ（ｓ）的例子
（＊）６—１７　　矩母函数的极限函数不是矩母函数的例子
第七章　　与大数定律、中心极限定理有关的反例
７—１　　随机变量序列｛ξｎ｝不服从于大数定律的例子
７—２　　随机变量序列｛ξｋ｝不满足马尔科夫条件，但服从大数定律的例子
７—３　　独立随机变量序列｛ξｋ｝不满足车贝谢夫大数定律的条件，但满足马尔科夫大数定律条件的例子
７—４　　独立随机变量序列｛ξｋ｝不满足格涅坚科大数定律的充要条件的例子
７—５　　马尔科夫条件满足，但柯尔莫哥洛夫强大数定律条件不满足的例子
７—６　　独立随机变量序列｛ξｋ｝不满足强大数定律的例子
７—７　　林德贝格条件不满足，但中心极限定理仍成立的例子
７—８　　费勒条件不满足，但中心极限定理仍成立的例子
７—９　　有关大数定律和中心极限定理之间关系的反例
７—１０　　随机变量序列｛ξｎ｝独立同分布，但不服从中心极限定理的例子
＊）７—１１　　Ｅ｜ξ｜２＝∞，Ｌｉｎｄｅｂｅｒｇ－Ｉéｖｙ中心极限定理仍成立的例子
（＊）７—１２　　｛ξｎｋ１≤ｋ≤ｎ，ｎ≥１｝满足ｕ，ａ，ｎ条件，但？｜ξｎｋ｜？０的例子
（＊）７—１３　　？∞，｛ξｎ｝仍服从强大数定律的例子
（＊）７—１４　　？（ξｎ－Ｅξｎ）ａ，ｓ收敛，但？Ｄξｎ＝∞的例子
第八章　　与充分统计量和完全统计量有关的反例
８—１　　并非一切统计量都是充分统计量的例子
（＊）８—２　　极小充分统计量不是完全充分统计量的例子
８—３　　次序统计量不是完全统计量的例子
８—４　　有界完全的分布族（或统计量）不是完全的分布族（或统计量）的例子
（＊）８—５　　可测函数ｆ（Ｘ）与有界充分完全统计量ｔ（Ｘ）独立，但ｆ（Ｘ）的分布与θ有关的例子
８—６　　充分统计量的函数不是充分统计量的例子

8－7　充分完全统计量的函数不是充分完全统计量的例子
8－8　指数族分布表达式中的T（X）不是充分完全统计量的例子
（*）8－9　在适当的条件下，当Ｔｉ（？ｉ）是θｉ（ｉ＝１，２）的充分统计量时，（Ｔ１（Ｘ１），Ｔ２（Ｘ２））是（θ１，θ２）的联合充分统计量的例子
第九章　与点估计有关的反例
9－1　参数不存在无偏估计的例子
9－2　无偏估计不是一致最小方差无偏估计的例子
9－3　Ｘ～｛ｆ（ｘ；θ），θ∈？｝，ｘ１，…，ｘｎ是其ｉｉｄ样本，参数θ的无偏估计是Ｘ１，Ｘ２，…，Ｘｎ的对称函数，但不是θ的ＵＭＶＵＥ的例子
（*）9－4　无偏估计存在，而ＵＭＶＵＥ不存在的例子
9－5　？是θ的无偏估计，而ｇ（？）不是ｇ（θ）的无偏估计的例子（其中ｇ（Ｘ）是Ｘ＝（Ｘ１，…，Ｘｎ）的Ｂｏｒｅｌ可测函数）
9－6　在一定的优良性准则下，用Ｓ？（Ｘｉ－？）２去估计σ２较之用Ｓ？×（Ｘｉ－？）２去估计σ２更优的例子
9－7　无偏估计不是一致（相合）估计的例子
9－8　无偏估计的方差低于Ｒａｏ－Ｃｒａｍéｒ不等式下界的例子
9－9　ＵＭＶＵＥ其方差达不到Ｒａｏ－Ｃｒａｍéｒ不等式下界的例子
9－10　充分统计量不是有效估计量的例子
（*）9－11　参数的ＵＭＶＵＥ不是参数的可容许估计的例子
9－12　某些条件不满足，但似然方程仍然存在一致（相合）解并且满足渐近正态性的例子
（*）9－13　参数θ的一致（相合）渐近正态估计？在θ点的渐近方差？中的ｖ（θ）低于Ｃ－Ｒ下界的例子
9－14　若Ｔｎ（Ｘ）是θ的一致估计，又｜Ｔｎ－θ｜≤Ａｎ＜∞，则Ｔｎ不是θ的均方一致估计〔即Ｅ（Ｔｎ－θ）２→０，（ｎ→∞）〕的例子
9－15　极大似然估计不是充分统计量的例子
9－16　极大似然估计不是有效估计的例子
9－17　似然方程的解不是极大似然估计的例子
（*）9－18　极大似然估计不是一致估计的例子
第十章　与假设检验有关的反例

１０—１　　不是单参数指数族，但具有单调似然比的例子
（＊）１０—２　　单边假设检验不存在一致最优势检验（记为ＵＭＰ检验）的例子
（＊）１０—３　　双边假设检验不存在ＵＭＰ检验的例子
（＊）１０—４　　检验函数φ对公共边界相似，但φ没有Ｎｅｙｍａｎ结构的例子
１０—５　　当正则条件不成立时，Ｗｉｌｋｓ定理中关于似然比极限分布的结论可以不成立的例子
１０—６　　对固定的样本大小而言，似然比检验可以不是无偏的例子
１０—７　　一个很坏的似然比检验的例子
第十一章　　与线性模型有关的反例
１１—１　　在非正态的条件下，参数的最小方差线性无偏估计不再是最小方差无偏估计的例子
第十二章　　与抽样理论有关的反例
１２—１　　当检查人员可能犯错误时，全面检查不一定比抽样检查更好的例子
主要参考文献